

Periodic-orbit theory of universal level correlations in quantum chaos

Sebastian Müller¹, Stefan Heusler², Alexander Altland³, Petr Braun^{4,5} and Fritz Haake⁴

¹*Department of Mathematics, University of Bristol, Bristol BS81TW, UK*

²*Institut für Didaktik der Physik, Universität Münster,
Wilhelm-Klemm Str. 10, 48149 Münster, Germany*

³*Institut für Theoretische Physik, Universität zu Köln, Zülpicher Str. 77, 50937 Köln, Germany*

⁴*Fachbereich Physik, Universität Duisburg-Essen, 47048 Duisburg, Germany and*

⁵*Institute of Physics, Saint-Petersburg University, 198504 Saint-Petersburg, Russia**

Using Gutzwiller's semiclassical periodic-orbit theory we demonstrate universal behavior of the two-point correlator of the density of levels for quantum systems whose classical limit is fully chaotic. We go beyond previous work in establishing the full correlator such that its Fourier transform, the spectral form factor, is determined for all times, below and above the Heisenberg time. We cover dynamics with and without time reversal invariance (from the orthogonal and unitary symmetry classes). A key step in our reasoning is to sum the periodic-orbit expansion in terms of a matrix integral, like the one known from the sigma model of random-matrix theory.

PACS numbers: 05.45.Mt, 03.65.Sq

I. INTRODUCTION

Classical chaos is characterized by sensitive dependence on initial conditions, and by the fact that long trajectories uniformly fill the available space. These classical features have profound consequences in quantum mechanics: Many quantum properties of chaotic systems become universal in the semiclassical limit, i.e., they no longer depend on the system in question but only on symmetries. For instance, as conjectured in [1] the statistics of energy eigenvalues is universal: Neighboring levels tend to repel each other, and the correlation function of the level density $\rho(E)$,

$$R(\epsilon) = \frac{1}{\bar{\rho}^2} \left\langle \rho \left(E + \frac{\epsilon}{2\pi\bar{\rho}} \right) \rho \left(E - \frac{\epsilon}{2\pi\bar{\rho}} \right) \right\rangle - 1, \quad (1)$$

(where the brackets and overbars denote averages over the center energy E) has the same form for all dynamics belonging to the same symmetry class. The most important classes are those of systems without any symmetries (unitary class; complex Hermitian Hamiltonians) and systems whose sole symmetry is time-reversal invariance (orthogonal class; real symmetric Hamiltonians).

If one accepts universal spectral correlations as a fact, a phenomenological prediction for $R(\epsilon)$ can be obtained by averaging over ensembles of systems sharing the same symmetries. Modelling the corresponding Hamiltonians through matrices, one arrives at random-matrix theory (RMT), as initially proposed by Wigner and Dyson in the fifties [2] in the context of atomic nuclei. RMT has since found broad applications in many fields of physics. For the correlation function $R(\epsilon)$, random-matrix averages yield the prediction [3, 4, 5]

$$R(\epsilon) = \begin{cases} -s(\epsilon)^2 & \text{unitary} \\ -s(\epsilon)^2 + s'(\epsilon) \left(\frac{1}{\pi} \text{Si}(\epsilon) - \frac{1}{2} \text{sgn}(\epsilon) \right) & \text{orthogonal} \end{cases} \quad (2)$$

with $s(\epsilon) = \frac{\sin \epsilon}{\epsilon}$ and $\text{Si}(\epsilon) = \int_0^\epsilon d\epsilon' s(\epsilon')$. It is convenient to access $R(\epsilon)$ through a complex correlation function $C(\epsilon)$ which is analytic in the energy upper half plane and connected with the real correlator as $R(\epsilon) = \lim_{\text{Im } \epsilon \rightarrow +0} \text{Re } C(\epsilon)$. The correlator $C(\epsilon)$ has the nice property of being retrievable by Borel summation [6] from its asymptotic expansion

$$C(\epsilon) \sim \sum_{n=2}^{\infty} (c_n + d_n e^{2i\epsilon}) \left(\frac{1}{\epsilon} \right)^n. \quad (3)$$

The foregoing power series contains both non-oscillatory terms and oscillatory terms proportional to $e^{2i\epsilon}$; the latter are responsible for the singularity (discontinuity at $\tau = 1$ of the first or third derivative in the unitary and orthogonal

*Electronic address: sebastian.muller@bristol.ac.uk, stefanheusler@gmx.de, alexal@thp.uni-koeln.de, petr.braun@uni-due.de, fritz.haake@uni-due.de

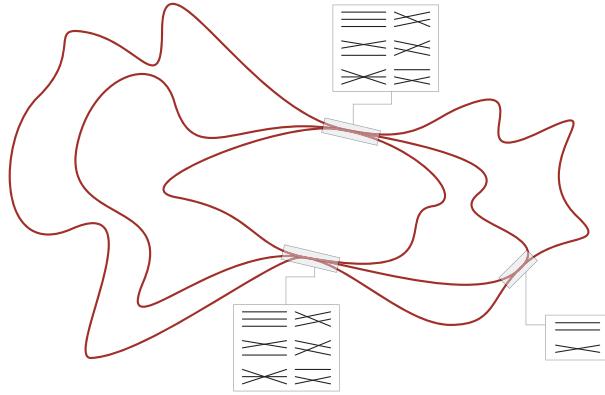


FIG. 1: Bunch of 72 (pseudo-)orbits differing in two 3-encounters and one 2-encounter (a pseudo-orbit is a set of several disjoined orbits; see below). Different orbits not resolved except in blowups of encounters

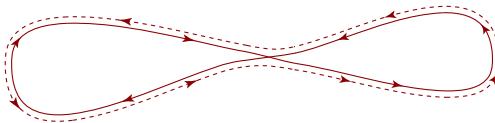


FIG. 2: Sieber-Richter pair

case, respectively) in the Fourier transform of the spectral correlator $K(\tau) = \frac{1}{\pi} \int_{-\infty}^{\infty} d\epsilon R(\epsilon) e^{2i\epsilon\tau}$ (the spectral form factor). For the unitary class we have $c_2 = -\frac{1}{2}$, $d_2 = \frac{1}{2}$ and all other coefficients vanish, whereas the orthogonal class involves $c_2 = -1$, $c_n = \frac{(n-3)!(n-1)}{2i^n}$ for $n \geq 3$, $d_2 = d_3 = 0$ and $d_n = \frac{(n-3)!(n-3)}{2i^n}$ for $n \geq 4$.

This phenomenological RMT result leaves open the question *why* chaotic systems behave universally, and that question we want to address in the present paper. Taking a semiclassical approach one can follow Gutzwiller [7] and express the level density as a sum over contributions of classical periodic orbits. The correlation function then turns into a sum over pairs of orbits. Systematic contributions to that double sum are due to pairs of orbits whose actions are sufficiently close for the associated quantum amplitudes to interfere constructively. The task of finding correlations in quantum spectra is translated into a classical one, namely to understand the correlations between actions of periodic orbits [8].

The first step in this direction was taken by Berry [9] who showed that pairs of *identical* or *mutually time-reversed* orbits explain the leading coefficient c_2 . Higher-order contributions c_n with $n \geq 3$ are due to a still not widely known feature of chaos, a certain “bunching” of periodic orbits [10]. A long orbit has many close “self-encounters” (see Fig. 1) where it comes close to itself in phase space (possibly up to time reversal, for dynamics with time reversal invariance). Such an orbit is just one of a closely packed “bunch”. All orbits in a bunch are nearly identical, except that the orbit pieces inside the encounters are “switched”.

Based on insight from disordered systems [11], the study of bunches of orbits differing in encounters was pioneered by Sieber and Richter [12] who derived the next-to-leading coefficient c_3 for time-reversal invariant systems from orbit pairs differing in just one encounter where two stretches of an orbit are close (see Fig. 2). Full agreement with all coefficients c_n from RMT was established by the present authors in [13, 14].

In none of these works oscillatory contributions could be obtained (Note however courageous forays by Keating [16] and Bogomolny and Keating [17]), due to the fact that Gutzwiller’s formula for the level density is divergent. To enforce convergence, one needs to allow for complex energies with imaginary parts large compared to the mean level spacing, $\text{Im } \epsilon \gg 1$. Oscillatory terms proportional to $e^{2i\epsilon}$ then become exponentially small and cannot be resolved within the conventional semiclassical approach.

In [18] we proposed a way around this difficulty that we here want to elaborate in detail. The key idea is to represent the correlation function through derivatives of a generating function involving spectral determinants. Two such representations are available and entail different semiclassical periodic-orbit expansions. One of them recovers the non-oscillatory part of the asymptotic expansion (3), essentially in equivalence to [12, 13, 14]; the other representation breaks new ground by giving the oscillatory part of (3).

The full random-matrix result also, and in fact most naturally, arises within an alternative semiclassical approxima-

tion scheme proposed by Berry and Keating in [19]. That scheme constrains the semiclassical periodic-orbit expansion of the spectral determinants $\det(E - H)$ to be real and to converge for real energy argument. Inserted into the generating function the resulting “Riemann-Siegel lookalike formula” for $\det(E - H)$ was shown in [20] to simultaneously produce both the non-oscillatory and oscillatory parts of the expansion (3). Interestingly, the Riemann-Siegel lookalike formula parallels an exact composition of the random-matrix averaged generating function originating from a Weyl group symmetry [21].

We want to stress that obtaining both the non-oscillatory and the oscillatory contributions to $R(\epsilon)$ also implies a derivation of level repulsion: By summing up the asymptotic series (3) we access the full $R(\epsilon)$ including its behavior for small ϵ , $R(\epsilon) \propto |\epsilon|^\beta$ (with $\beta = 2$ for the unitary and $\beta = 1$ for the orthogonal class), indicating a suppression of small energy differences. To get access to small ϵ the oscillatory contributions (varying on the scale of the mean level spacing) are of crucial importance.

For our semiclassical approach to spectral statistics we have drawn inspiration from a field-theoretical implementation of RMT, the nonlinear sigma model. There, the ensemble averaged generating function is ultimately transformed into an integral over matrices. In a stationary-phase approximation to that integral, two different saddles are found to respectively contribute non-oscillatory and oscillatory terms. The relation between these two contributions bears close resemblance to semiclassics. Moreover the topologically different structures of orbit bunches in semiclassics turn out to be analogous to Feynman diagrams in a perturbative treatment of the sigma model, with encounters corresponding to vertices. That analogy helps to do the sums required in semiclassics and to compact the semiclassical series to a matrix integral like the one in the sigma model. This derivation has to be contrasted with a previous field-theoretical approach to spectral universality, the ballistic sigma model [22, 23].

The present paper is organized as follows. We start by explaining the generating-function approach (Section II), and give the pertinent generalization of Berry’s diagonal approximation (Section III). We then evaluate the contributions of bunches of periodic orbits to the generating function and the correlation function (Section IV) and sum them up (Section V). Our reasoning culminates in the semiclassical construction of a sigma model lookalike (Section VI). For pedagogical reasons, Sections III-VI are devoted to the unitary symmetry class, the simplest case. The generalizations necessary for time-reversal invariant dynamics are listed in Section VII.

The Appendix A contains a derivation of contraction rules needed for the semiclassical construction of the matrix integral. Further appendices contain (B) a brief introduction to the supersymmetric sigma model, (C, D) two replica versions of the sigma model and their relation to the semiclassical theory, (E) a characterization of orbit topologies in terms of permutations, and (F) the energy average employed in the definition of the correlator and the connection between the real and complex correlation functions.

II. GENERATING FUNCTION

A. Definition

All approaches to resolve non-oscillatory and oscillatory contributions start with representing the correlation function through derivatives of a generating function. To get there we express the complex correlator $C(\epsilon)$ in terms of Green functions $(E - H)^{-1}$ of the Hamiltonian H ,

$$C(\epsilon^+) = \frac{1}{2\pi^2 \bar{\rho}^2} \left\langle \text{tr} \left(E + \frac{\epsilon^+}{2\pi\bar{\rho}} - H \right)^{-1} \text{tr} \left(E - \frac{\epsilon^+}{2\pi\bar{\rho}} - H \right)^{-1} \right\rangle - \frac{1}{2}, \quad (4)$$

where the superscript + denotes a positive imaginary part $i\eta$. The average will be understood as over the real center energy E , with an averaging window large compared to the mean level spacing (see Appendix F). The first Green function in (4) is a retarded one due to the positive imaginary part of the energy argument while the second is an advanced one.

A suitable generating function involves four spectral determinants $\Delta(E) = \det(E - H)$,

$$Z(\epsilon_A, \epsilon_B, \epsilon_C, \epsilon_D) = \left\langle \frac{\Delta \left(E + \frac{\epsilon_C}{2\pi\bar{\rho}} \right) \Delta \left(E - \frac{\epsilon_D}{2\pi\bar{\rho}} \right)}{\Delta \left(E + \frac{\epsilon_A}{2\pi\bar{\rho}} \right) \Delta \left(E - \frac{\epsilon_B}{2\pi\bar{\rho}} \right)} \right\rangle. \quad (5)$$

The four parameters $\epsilon_A, \epsilon_B, \epsilon_C, \epsilon_D$ denote offsets from the center energy E , made dimensionless by referral to the

mean level spacing $1/\bar{\rho}$; they are all taken to include a positive imaginary part $i\eta$ ¹. It is worth noting the symmetry

$$Z(\epsilon_A, \epsilon_B, \epsilon_C, \epsilon_D) = Z(\epsilon_A, \epsilon_B, -\epsilon_D, -\epsilon_C) \quad (6)$$

which in the framework of the sigma-model treatment of RMT reflects a Weyl symmetry [21]. The normalization $Z(\epsilon_A, \epsilon_B, \epsilon_A, \epsilon_B) = 1$ also deserves being mentioned.

The correlation function can be accessed from Z in two different ways. In the standard “columnwise” representation, we take two derivatives w.r.t. ϵ_A and ϵ_B , and subsequently identify all energy increments,

$$\begin{aligned} -2 \frac{\partial^2 Z}{\partial \epsilon_A \partial \epsilon_B} \Big|_{(\parallel)} &= \frac{1}{2(\pi\bar{\rho})^2} \left\langle \text{tr} \left(E + \frac{\epsilon_A}{2\pi\bar{\rho}} - H \right)^{-1} \text{tr} \left(E - \frac{\epsilon_B}{2\pi\bar{\rho}} - H \right)^{-1} \frac{\Delta(E + \frac{\epsilon_C}{2\pi\bar{\rho}}) \Delta(E - \frac{\epsilon_D}{2\pi\bar{\rho}})}{\Delta(E + \frac{\epsilon_A}{2\pi\bar{\rho}}) \Delta(E - \frac{\epsilon_B}{2\pi\bar{\rho}})} \right\rangle \Big|_{(\parallel)} \\ &= \frac{1}{2(\pi\bar{\rho})^2} \left\langle \text{tr} \left(E + \frac{\epsilon^+}{2\pi\bar{\rho}} - H \right)^{-1} \text{tr} \left(E - \frac{\epsilon^+}{2\pi\bar{\rho}} - H \right)^{-1} \right\rangle = C(\epsilon^+) + \frac{1}{2}; \end{aligned} \quad (7)$$

$$(\parallel) : \epsilon_A = \epsilon_B = \epsilon_C = \epsilon_D = \epsilon^+. \quad (8)$$

Here we used the identity $\det = \exp \text{tr} \ln$. The notation (\parallel) signals that when equating $\epsilon_A, \epsilon_B, \epsilon_C$ and ϵ_D and accounting for the signs in (5), we have identified the four energy arguments in (5) in a “columnwise” way.

An alternative representation of the correlation function can be obtained if we identify the energy arguments in a “crosswise” way (in view of the signs in (5)),

$$\begin{aligned} -2 \frac{\partial^2 Z}{\partial \epsilon_A \partial \epsilon_B} \Big|_{(\times)} &= \frac{1}{2(\pi\bar{\rho})^2} \left\langle \text{tr} \left(E + \frac{\epsilon^+}{2\pi\bar{\rho}} - H \right)^{-1} \text{tr} \left(E - \frac{\epsilon^+}{2\pi\bar{\rho}} - H \right)^{-1} \frac{\Delta \left(E - \frac{\epsilon^-}{2\pi\bar{\rho}} \right) \Delta \left(E + \frac{\epsilon^-}{2\pi\bar{\rho}} \right)}{\Delta \left(E + \frac{\epsilon^+}{2\pi\bar{\rho}} \right) \Delta \left(E - \frac{\epsilon^+}{2\pi\bar{\rho}} \right)} \right\rangle \\ (\times) : \quad \epsilon_A &= \epsilon_B = \epsilon^+, \epsilon_C = \epsilon_D = -\epsilon^- \equiv -\epsilon + i\eta. \end{aligned} \quad (9)$$

In (9), the ratio of determinants² becomes unity only in the limit $\eta \rightarrow 0$, i.e., one can access $R(\epsilon)$ as

$$\text{Re} \lim_{\eta \rightarrow 0} \left(-2 \frac{\partial^2 Z}{\partial \epsilon_A \partial \epsilon_B} \Big|_{(\times)} \right) = R(\epsilon) + \frac{1}{2}. \quad (10)$$

B. Periodic-orbit representation

We are now prepared to semiclassically approximate the generating function in terms of classical periodic orbits. Our starting point is Gutzwiller’s formula for the trace of the Green function,

$$\text{tr}(E^+ - H)^{-1} \sim -i\pi\bar{\rho}(E^+) - \frac{i}{\hbar} \sum_a T_a F_a e^{iS_a(E^+)/\hbar}. \quad (11)$$

Herein, the first term involves the average level density $\bar{\rho}$ given by Weyl’s law as $\bar{\rho} = \frac{\Omega}{(2\pi\hbar)^2}$ (for systems with two degrees of freedom), with Ω the volume of the energy shell. The orbit a is represented by its period T_a , action S_a , and stability amplitude F_a (defined to include the Maslov phase factor but not the period)

³

¹ Due to the imposed positivity of the imaginary parts of the energy offsets the determinants involving ϵ_A and ϵ_C are associated with retarded Green functions. The minus signs in front of ϵ_B and ϵ_D in (5) let the respective spectral determinants be associated with advanced Green functions. Note that the present sign convention differs from the one in our letter [18].

² The above ratio averages to $Z(\epsilon^+, \epsilon^+, -\epsilon^-, -\epsilon^-)$ which can also be written as $Z(\epsilon^+, \epsilon^+, -\epsilon^-, -\epsilon^-) = \langle \exp[2\pi i \Delta N_\eta] \rangle$ with $\Delta N_\eta = N_\eta(E + \frac{\epsilon}{2\pi\bar{\rho}}) - N_\eta(E - \frac{\epsilon}{2\pi\bar{\rho}})$. Here N_η denotes the level staircase smeared out over a range η , $N_\eta(E) = \sum_k \theta_\eta(E - E_k)$ with $\theta_\eta(x) = \pi^{-1} \text{Im} \ln(-x + i\eta)$ the similarly smeared step function. The function $Z(\epsilon^+, \epsilon^+, -\epsilon^-, -\epsilon^-)$ could be used as a starting point for calculating the oscillatory part of the correlator $C(\epsilon)$, and then some similarity with the courageous foray of Bogomolny and Keating [17] mentioned in the Introduction would arise. Indeed, replacing N_η by the sum of its energy average and a fluctuating part, $N_\eta = \langle N_\eta \rangle + N_\eta^{\text{fluct}}$, and treating N_η^{fluct} in a Gaussian approximation we get $\langle \exp[2\pi i \Delta N_\eta] \rangle \sim e^{2i\epsilon} \exp[-2\pi^2 \langle (\Delta N_\eta^{\text{fluct}})^2 \rangle]$. The second factor contains the number variance in the exponent; the first factor signals that the oscillatory part of the two-point correlator becomes accessible. This rough argument anticipates the diagonal approximation below.

³ We disregard orbits which are multiple repetitions of a shorter orbit since these are exponentially suppressed in the limit of long periods. As regards orbits incorporating multiple *approximate* traversals of short orbits, the respective corrections were shown to be immaterial for the non-oscillatory part of the spectral correlator (see [14]) and can similarly be shown to cancel in our four-determinant generating function Z . Interestingly, such orbits do contribute to the averaged product of *two* spectral determinants [15].

To proceed to a semiclassical approximation of the spectral determinant, we integrate and exponentiate in (11) as $\Delta(E^+) \propto \exp\left(\int^{E^+} dE' \text{tr} \frac{1}{E' - H}\right)$. The resulting exponent inherits the smoothed number $\overline{N}(E)$ of eigenvalues below E and an orbit sum from the trace formula (11),

$$\Delta(E^+) \propto \exp\left(-i\pi\overline{N}(E^+) - \sum_a F_a e^{iS_a(E^+)/\hbar}\right) = e^{-i\pi\overline{N}(E^+)} \sum_A F_A (-1)^{n_A} e^{iS_A(E^+)/\hbar}. \quad (12)$$

The last member of the foregoing chain of equations has the exponentiated orbit sum expanded into a sum over *pseudo-orbits*, i. e., non-ordered sets of periodic orbits A [19]. In that pseudo-orbit sum, n_A is the number of orbits inside A , S_A the sum of their actions, and F_A the product of their stability amplitudes F_a . The sum over A includes the empty set, whose contribution is unity. (Note that in the foregoing expressions we suppressed a proportionality constant due to the lower limit of the E' -integral. That factor is irrelevant since it cancels in the generating function.)

Eq. (12) applies to energies with a positive imaginary part. For energies with a negative imaginary part complex conjugation yields

$$\begin{aligned} \Delta(E^-) &= \Delta(E^+)^* \propto \exp\left(i\pi\overline{N}(E^-) - \sum_a F_a^* e^{-iS_a(E^-)/\hbar}\right) \\ &= e^{i\pi\overline{N}(E^-)} \sum_A F_A^* (-1)^{n_A} e^{-iS_A(E^-)/\hbar}. \end{aligned} \quad (13)$$

For inverse spectral determinants, the minus sign in the periodic-orbit sum and hence the factor $(-1)^{n_A}$ in the pseudo-orbit sum disappear,

$$\begin{aligned} \Delta(E^+)^{-1} &= (\Delta(E^-)^{-1})^* \propto \exp\left(i\pi\overline{N}(E^+) + \sum_a F_a e^{iS_a(E^+)/\hbar}\right) \\ &= e^{i\pi\overline{N}(E^+)} \sum_A F_A e^{iS_A(E^+)/\hbar}. \end{aligned} \quad (14)$$

The four spectral determinants in Z now yield semiclassical asymptotics as a fourfold sum over pseudo-orbits,

$$\begin{aligned} Z \sim Z^{(1)} &= \left\langle e^{i\pi\overline{N}[E+\epsilon_A/(2\pi\bar{\rho})]} \sum_A F_A e^{iS_A[E+\epsilon_A/(2\pi\bar{\rho})]/\hbar} \right. \\ &\quad \times e^{-i\pi\overline{N}[E-\epsilon_B/(2\pi\bar{\rho})]} \sum_B F_B^* e^{-iS_B[E-\epsilon_B/(2\pi\bar{\rho})]/\hbar} \\ &\quad \times e^{-i\pi\overline{N}[E+\epsilon_C/(2\pi\bar{\rho})]} \sum_C F_C (-1)^{n_C} e^{iS_C[E+\epsilon_C/(2\pi\bar{\rho})]/\hbar} \\ &\quad \left. \times e^{i\pi\overline{N}[E-\epsilon_D/(2\pi\bar{\rho})]} \sum_D F_D^* (-1)^{n_D} e^{-iS_D[E-\epsilon_D/(2\pi\bar{\rho})]/\hbar} \right\rangle, \end{aligned} \quad (15)$$

where all offset variables $\epsilon_A, \epsilon_B, \epsilon_C, \epsilon_D$ have positive imaginary parts η . Here the superscript 1 signals that (15) is a semiclassical approximation based on the Gutzwiller trace formula which requires

$$\eta \gg 1 \quad (16)$$

for convergence such that terms vanishing in that limit (such as terms of the order $e^{-\eta}$) will be lost. The complement missed by $Z^{(1)}$ becomes important for real energies (a manifestation of the Stokes phenomenon of asymptotic analysis) and will be introduced below.

Upon expanding as $\overline{N}(E \pm \frac{\epsilon}{2\pi\bar{\rho}}) \sim \overline{N}(E) \pm \frac{\epsilon}{2\pi}$ and $S(E \pm \frac{\epsilon}{2\pi\bar{\rho}})/\hbar \sim S(E) \pm \frac{T\epsilon}{2\pi\hbar\bar{\rho}} = S(E) \pm T\epsilon/T_H$ where $T_H = 2\pi\hbar\bar{\rho} = \frac{\Omega}{2\pi\hbar}$ is the Heisenberg time, we simplify the quantity $Z^{(1)}$ to

$$\begin{aligned} Z^{(1)} &= e^{i(\epsilon_A + \epsilon_B - \epsilon_C - \epsilon_D)/2} \times \left\langle \sum_{A,B,C,D} F_A F_B^* F_C F_D^* (-1)^{n_C + n_D} \right. \\ &\quad \left. \times e^{i(S_A(E) - S_B(E) + S_C(E) - S_D(E))/\hbar} e^{i(T_A \epsilon_A + T_B \epsilon_B + T_C \epsilon_C + T_D \epsilon_D)/T_H} \right\rangle. \end{aligned} \quad (17)$$

The foregoing sum over pseudo-orbit quadruplets combines classical chaos with quantum interference due to the phase factor $e^{i\Delta S/\hbar} \equiv e^{i(S_A(E) - S_B(E) + S_C(E) - S_D(E))/\hbar}$. For most quadruplets the interference is destructive since for

$\hbar \rightarrow 0$ the phase factor $e^{i\Delta S/\hbar}$ oscillates rapidly as the energy varies and therefore vanishes after averaging. For constructive interference, the action mismatch ΔS must be small, at most of the order of \hbar ; to within such a quantum mismatch the cumulative action of the pseudo-orbits $B \cup D$ must coincide with the cumulative action of $A \cup C$. Such action matching happens systematically only if the orbits in $B \cup D$ either coincide with those in $A \cup C$ (and thus constitute the “diagonal” contributions, see Sect. III) or differ from them only by their connections inside encounters (“off-diagonal” contributions) as in Figs. 1,2. Indeed, then, the phenomenon of orbit bunching is crucial for spectral universality.

Now imagine $Z^{(1)}$ calculated from (17) and then analytically continued to real energy offsets, $\eta \rightarrow 0^+$. If instead of $Z^{(1)}$ we had the full generating function we could obtain both the non-oscillatory and the oscillatory contributions to $C(\epsilon)$ by taking derivatives w.r.t. ϵ_A and ϵ_B and identifying the energy increments either in the columnwise way (7) or the crosswise one (9). However, for $Z^{(1)}$ the columnwise procedure (7) cannot yield oscillatory terms since it makes the (Weyl) exponential $e^{i(\epsilon_A + \epsilon_B - \epsilon_C - \epsilon_D)/2}$ collapse to unity. In contrast, with the crosswise representation (9) the signs of ϵ_C and ϵ_D are flipped and the Weyl factor turns into $e^{2i\epsilon}$. Hence the crosswise representation recovers only the oscillatory Fourier component proportional to $e^{2i\epsilon}$.

The random-matrix result suggests that to get the *full correlation function* one has to add both results

$$R(\epsilon) = -\text{Re} \left\{ 2 \frac{\partial^2 Z^{(1)}}{\partial \epsilon_A \partial \epsilon_B} \Big|_{(\parallel)} + 2 \frac{\partial^2 Z^{(1)}}{\partial \epsilon_A \partial \epsilon_B} \Big|_{(\times)} + \frac{1}{2} \right\}. \quad (18)$$

Here the first and the third summand are due to (7) while the second summand is due to (10)⁴. Of course it is vital to justify this additivity independently from RMT, a task to be attacked below.

The “crosswise” term $\frac{\partial^2 Z^{(1)}}{\partial \epsilon_A \partial \epsilon_B} \Big|_{(\times)}$ differs from the “columnwise” term $\frac{\partial^2 Z^{(1)}}{\partial \epsilon_A \partial \epsilon_B} \Big|_{(\parallel)}$ because ϵ_C and ϵ_D are interchanged and flipped in sign. This interchange and sign flip may as well be performed on the level of the generating function, i.e., we can write

$$\frac{\partial^2 Z^{(1)}}{\partial \epsilon_A \partial \epsilon_B} \Big|_{(\times)} = \frac{\partial^2 Z^{(2)}}{\partial \epsilon_A \partial \epsilon_B} \Big|_{(\parallel)} \quad (19)$$

$$Z^{(2)}(\epsilon_A, \epsilon_B, \epsilon_C, \epsilon_D) = Z^{(1)}(\epsilon_A, \epsilon_B, -\epsilon_D, -\epsilon_C) \quad (20)$$

Instead of adding the crosswise representation in (18) we can use $Z^{(1)} + Z^{(2)}$ as a generating function, and get $R(\epsilon)$ in full through the columnwise (or crosswise) representation alone. Therefore no extra work is needed for the oscillatory part of the two-point correlation function, once the “standard” part $Z^{(1)}$ of the asymptotics of Z is found. The resulting asymptotics of the generating function

$$Z \sim Z^{(1)} + Z^{(2)} \quad (21)$$

respects the Weyl symmetry and remains valid when the arguments are real ($\eta = 0$). We note that if one wants to evaluate $Z^{(2)}$ using the Gutzwiller trace formula, the sign flip in ϵ_C and ϵ_D means that the corresponding arguments of $Z^{(2)}$ should be taken with large *negative* imaginary parts.

C. Riemann-Siegel lookalike

As pointed out by Keating and Müller in [20], the additive structure (21) of the generating function most naturally arises with the help of the Riemann-Siegel lookalike formula [19].

The essential idea is to eliminate a drawback of the semiclassical approximation (12,13) of the spectral determinant, namely its failure to manifestly preserve the reality of the energy eigenvalues, i. e., the unitarity of the quantum dynamics. In the limit $\eta \rightarrow 0$, $\Delta(E^+)$ and $\Delta(E^-)$ must become real and identical to each other. That property is not manifest in but may be imposed on Eqs. (12,13). By imposing this property, it was shown in [19] that a duality arises between “long” pseudo-orbits (sum of periods of contributing orbits larger than half the Heisenberg time T_H) and “short” ones (cumulative period smaller than $T_H/2$); the overall contribution of long orbits is the

⁴ Critical readers may miss the $\frac{1}{2}$ from (10). This summand does not show up here since it is non-oscillatory and cannot be retrieved from the crosswise representation.

complex conjugate of the contribution arising from short orbits. The ensuing Riemann-Siegel lookalike expresses the real spectral determinant $\Delta(E)$ as a sum over pseudo-orbits with durations shorter than $T_H/2$,

$$\Delta(E) \propto e^{-i\pi\bar{N}(E)} \sum_{A (T_A < T_H/2)} F_A (-1)^{n_A} e^{iS_A/\hbar} + \text{c.c.} \quad (22)$$

To evaluate the generating function Z one now uses the Riemann-Siegel formula for the two spectral determinants in the numerator. For the determinants in the denominator we are not allowed to use Riemann-Siegel as the non-zero imaginary parts with appropriate signs in ϵ_A, ϵ_B are crucial for the definition of Z , and we just stick to Eq. (14). Since each determinant in the numerator becomes the sum of two mutually complex conjugate terms, Z has altogether four summands which involve the four phases $\pm i\pi [\bar{N}(E + \epsilon_C/(2\pi\bar{\rho})) - \bar{N}(E - \epsilon_D/(2\pi\bar{\rho}))]$ and $\pm i\pi [\bar{N}(E + \epsilon_C/(2\pi\bar{\rho})) + \bar{N}(E - \epsilon_D/(2\pi\bar{\rho}))]$. The two latter contributions oscillate rapidly as functions of E and thus vanish after averaging.

The two surviving additive terms yield the high-energy asymptotics of the generating function $Z \sim Z^{(1)} + Z^{(2)}$ found in the previous subsection, according to (17, 20), except that the sums over C and D are now restricted to pseudo-orbits with durations shorter than $T_H/2$. To make that difference irrelevant it is convenient to evaluate the pseudo-orbit sum for large imaginary parts η of the dimensionless energy offsets in the arguments of the action, $S_A(E) \rightarrow S_A(E + i\eta/2\pi\bar{\rho})$, see (16). That artifice makes sure that in the relevant integrals over orbit periods to be encountered below the integrand falls to exponentially small magnitude as the integration variable rises to near half the Heisenberg time. If all four energy offsets $\epsilon_{A,B,C,D}$ are thus provided with large positive imaginary parts, the two additive pieces obtained through the Riemann-Siegel lookalike become respectively $Z^{(1)}(\epsilon_A, \epsilon_B, \epsilon_C, \epsilon_D)$ and $Z^{(2)}(\epsilon_A, \epsilon_B, \epsilon_C^*, \epsilon_D^*) = Z^{(1)}(\epsilon_A, \epsilon_B, -\epsilon_D^*, -\epsilon_C^*)$. Complex conjugation of the last two arguments of $Z^{(2)}$ is needed to enable its calculation in terms of periodic orbits according to (17). At the final stage we use analyticity to let η go to zero; the sum of $Z^{(1)}$ and $Z^{(2)}$ yields then the above asymptotic representation (21).

To obtain a quantitative understanding of spectral statistics, we now have to evaluate the quadruple sum over pseudo-orbits in $Z^{(1)}$, taking into account orbits in $B \cup D$ that either (diagonal contributions) coincide with or (off-diagonal contributions) at most differ in encounters from the orbits in $A \cup C$.

III. DIAGONAL APPROXIMATION

In the diagonal approximation we effectively neglect all correlations between orbits that are not identical. If we represent the generating functions as a sum over A, B, C, D as in (17) this means we should keep only contributions where the orbits in $A \cup C$ are repeated inside $B \cup D$. However we can save labor if we write the four spectral determinants entering our generating function like the second member in (12), i. e., with the help of exponentiated orbit sums. We thus obtain the equivalent representation of the generating function

$$Z^{(1)} = e^{i(\epsilon_A + \epsilon_B - \epsilon_C - \epsilon_D)/2} \left\langle \prod_a z_a \right\rangle \quad (23)$$

wherein each periodic orbit participates with the factor

$$z_a = \exp \left[F_a e^{\frac{i}{\hbar} S_a(E)} \underbrace{\left(e^{i\frac{T_a}{T_H}\epsilon_A} - e^{i\frac{T_a}{T_H}\epsilon_C} \right)}_{\equiv f_{AC}^a} + F_a^* e^{-\frac{i}{\hbar} S_a(E)} \underbrace{\left(e^{i\frac{T_a}{T_H}\epsilon_B} - e^{i\frac{T_a}{T_H}\epsilon_D} \right)}_{\equiv f_{BD}^a} \right] \quad (24)$$

If we assume that contributions of different periodic orbits are uncorrelated, $Z^{(1)}$ factorizes into a product of energy-averaged orbit factors z_a ,

$$Z_{\text{diag}}^{(1)} = e^{i(\epsilon_A + \epsilon_B - \epsilon_C - \epsilon_D)/2} \prod_a \langle z_a \rangle. \quad (25)$$

Inasmuch as the stability coefficient $F_a \sim e^{-\lambda T_a/2}$ is small for long orbits we may expand the exponential in z_a and limit ourselves to the first nontrivial term, $\langle z_a \rangle \sim 1 + |F_a|^2 f_{AC}^a f_{BD}^a \approx e^{|F_a|^2 f_{AC}^a f_{BD}^a}$; here terms proportional to $e^{\pm i S(E)/\hbar}$ and $e^{\pm 2i S(E)/\hbar}$ have been eliminated by the energy average. We so obtain

$$Z_{\text{diag}}^{(1)} = e^{i(\epsilon_A + \epsilon_B - \epsilon_C - \epsilon_D)/2} \exp \left(\sum_a |F_a|^2 f_{AC}^a f_{BD}^a \right). \quad (26)$$

This result allows for an intuitive interpretation: If we assume that the non-identical orbits are uncorrelated the exponent in the product of all orbit factors z_a , namely $\Phi = \sum_a [F_a e^{iS_a(E)/\hbar} f_{AC}^a + F_a^* e^{-iS_a(E)/\hbar} f_{BD}^a]$, effectively becomes a sum of many independent random contributions; due to the central limit theorem Φ must then obey Gaussian statistics such that $\langle e^\Phi \rangle = e^{\langle \Phi^2 \rangle/2}$ which just yields (26).

To bring our result to a more explicit final form, we need to invoke ergodicity in the form of the sum rule of Hannay and Ozorio de Almeida [3, 24]: To sum a smooth function of the period T over an ensemble of orbits weighed with $|F_a|^2$, we may as well compute an integral over periods weighed with $\frac{1}{T}$,

$$\sum_a |F_a|^2(\dots) \approx \int_{T_0}^{\infty} \frac{dT}{T}(\dots); \quad (27)$$

here T_0 is some minimal period starting from which orbits behave ergodically. Employing the sum rule (27) we get

$$\sum_a |F_a|^2 f_{AC}^a f_{BD}^a \sim \int_{T_0}^{\infty} \frac{dT}{T} \left(e^{i\frac{T}{T_H}\epsilon_A} - e^{i\frac{T}{T_H}\epsilon_C} \right) \left(e^{i\frac{T}{T_H}\epsilon_B} - e^{i\frac{T}{T_H}\epsilon_D} \right).$$

Since for $T \ll T_H$ the integrand is $\propto (T/T_H)^2$ we can set the lower integration limit to 0, accepting an error of the order $(T_0/T_H)^2$. Then using $\int_0^{\infty} \frac{dx}{x} (e^{iax} - e^{ibx}) = \ln \frac{b}{a}$ we arrive at the diagonal part of the generating function

$$Z_{\text{diag}}^{(1)} = e^{\frac{i}{2}(\epsilon_A + \epsilon_B - \epsilon_C - \epsilon_D)} \frac{(\epsilon_A + \epsilon_D)(\epsilon_C + \epsilon_B)}{(\epsilon_A + \epsilon_B)(\epsilon_C + \epsilon_D)}. \quad (28)$$

The diagonal approximation yields an estimate for the correlation function $C(\epsilon)$. Taking into account both $Z_{\text{diag}}^{(1)}$ and its counterpart $Z_{\text{diag}}^{(2)}$ defined in Eq. (20) we obtain

$$C_{\text{diag}}(\epsilon) = -2 \frac{\partial^2 (Z_{\text{diag}}^{(1)} + Z_{\text{diag}}^{(2)})}{\partial \epsilon_A \partial \epsilon_B} \Big|_{(\parallel)} - \frac{1}{2} = \frac{1}{2(i\epsilon)^2} - \frac{e^{2i\epsilon}}{2(i\epsilon)^2}. \quad (29)$$

This result already reproduces the full RMT prediction for the unitary symmetry class under discussion. It remains to explain, within our semiclassical approach, why all off-diagonal contributions do in fact vanish.

IV. OFF-DIAGONAL CONTRIBUTIONS

Off-diagonal contributions arise from quadruplets of pseudo-orbits where not all orbits of $A \cup C$ are simply repeated inside $B \cup D$; some orbits of $A \cup C$ and $B \cup D$ differ in encounters. By studying these encounters in detail we shall show that for systems without time-reversal invariance the off-diagonal contributions cancel mutually in a nontrivial way. This also sets the stage for a generalization to time-reversal invariant dynamics where this cancellation does not occur, and off-diagonal quadruplets of pseudo-orbits give rise to terms of higher order in $1/\epsilon$.

A. Encounters and orbit bunches

Long periodic orbits often come close to themselves, and one another, in phase space. We speak of an l -encounter whenever l stretches of one or more periodic orbits in a chaotic system approach in this way. Each stretch begins and ends at phase-space points to be called ‘‘entrance ports’’ and ‘‘exit ports’’. These ports can be uniquely defined by setting an upper bound for the separation between the encountering stretches. The parts of an orbit in between encounter stretches will be called ‘‘links’’.

Now the crucial point is that the orbits and pseudo-orbits in hyperbolic motion exist in bunches. The members of a bunch differ noticeably only in encounters where the ports are connected differently by encounter stretches. In contrast the links are very nearly identical. All members of such a bunch have almost identical actions. If two of them are taken as $A \cup C$ and $B \cup D$ in Eq. (17) their contributions interfere constructively and need to be taken into account for a full understanding of spectral statistics.

The phenomenon of bunching is a consequence of hyperbolicity. For a system with two degrees of freedom⁵, hyperbolicity means that deflections between phase-space points in the three-dimensional energy shell can be split

⁵ In the present paper we shall stick to two-dimensional systems for convenience but the generalization to systems with arbitrarily many degrees of freedom along the lines of [14, 25] is straightforward.

into three components: one pointing along the flow, one unstable component u , and one stable component s . As two close-by phase-space points move, the unstable component of their separation grows exponentially whereas the stable component shrinks exponentially. In the limit of large times the rate of growth or shrinking is given by the Lyapunov exponent λ as

$$u(t) \sim u(0)e^{\lambda t}, \quad s(t) \sim s(0)e^{-\lambda t}. \quad (30)$$

As a consequence, given any two encounter stretches that are long compared with the characteristic size of the system and close enough to linearize the motion as in (30) we can draw a piece of a physical trajectory that starts close to the first stretch and ends close to the second stretch. Its separation from the first stretch must point along the unstable direction at the respective phase-space points (such that it approaches the first stretch for large negative times) whereas the separation from the second stretch must point along the stable direction (such that it approaches that stretch for large positive times). The new piece of trajectory is thus located where the unstable manifold of one encounter stretch of $A \cup C$ intersects the stable manifold of the other stretch [26]. As a consequence we can reconnect the ports of each encounter in any way we like while leaving the links in between the encounter stretches practically unchanged. (Note that subsequent to the construction just explained exponentially small corrections to the encounter stretches and links are needed to connect them to continuous orbits.)

B. Structures of pseudo-orbit quadruplets

The pseudo-orbit quadruplets can have many different topologies some of which are shown in Fig. 3. For their classification we introduce the notion of a *structure* by which we mean a diagram like the ones in Fig. 3 but with the encounters and their ports numbered and each periodic orbit assigned to a particular pseudo-orbit. All structures can be generated if we proceed as follows:

- S1: *Encounters* – We consider any number V of encounters. For each of the encounters $\sigma = 1, 2, \dots, V$, we take into account all possible numbers of stretches $l(\sigma) = 2, 3, 4, \dots$. The $l(\sigma)$ entrance and exit ports of each encounter are numbered from 1 to $l(\sigma)$ in such a way that the i -th entrance port is connected to the i -th exit port before reconnection, and to the $(i-1)$ -st exit port after reconnection; the first entrance reconnects to the last exit. The encounters can thus be depicted as in Fig. 4 where full and dashed lines denote the encounter stretches before and after the reconnection.
- S2: *Links* – We connect the ports by links in all possible ways. Each link must lead from an exit port to an entrance port of either the same or a different encounter.
- S3: *Orbits* – We include all possible distributions of the orbits among the pseudo-orbits A, B, C, D . The graph formed by all links and encounter stretches before reconnection falls into a certain number of disjoint parts each standing for a periodic orbit. Each of these can be included in either A or C . After changing connections inside the encounters the graph consists of a (possibly different) number of disjoint orbits which have to be distributed between B and D .

Thus equipped we shall split the sum over quadruplets in (17) into a sum over structures and a sum over quadruplets associated to each structure. The sum over quadruplets associated to each structure will be done using ergodicity. We shall determine the likelihood for having encounters inside and between periodic orbits; this likelihood depends on the length of the links and the phase-space separation between the encounter stretches. The sum can then be replaced by suitable integrals. The subsequent summation over structures becomes a purely combinatorial problem.

When doing the sums and integrals mentioned, we have to be aware of a certain overcounting. Since our definition of a structure involves an ordering of encounters and ports, each physical quadruplet can be associated with a structure in several different ways. To begin with, the encounters of a given quadruplet can be numbered in any one of $V!$ different orders. Then, in each encounter any of its $l(\sigma)$ entrance ports can be chosen as the first one. Note that as soon as the first entrance in an encounter is chosen, the numbers of all other encounter ports of the quadruplet become uniquely defined by our numbering scheme. In all, we get $V! \prod_{\sigma} l(\sigma)$ alternatives either leading to different structures or to the same structure but with different numberings and thus to a different set of the integration parameters. In both cases after summation over structures and integration over the continuous parameters the same quadruplet is taken into account several times. In order to compensate this overcounting, the contribution of each structure needs to be divided by $V! \prod_{\sigma} l(\sigma)$. Details can be found in Appendix E where a formal definition of structures in terms of permutations is given.

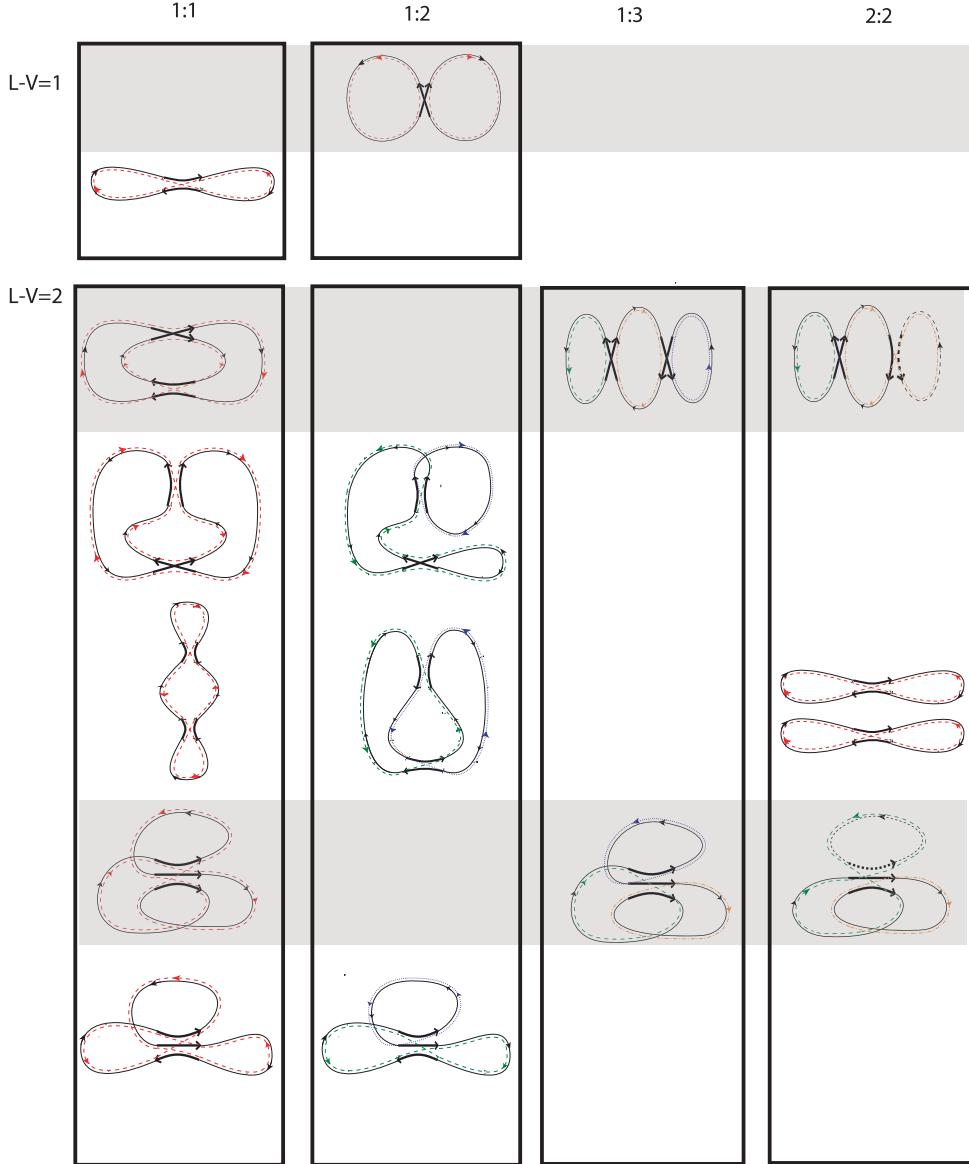


FIG. 3: Physically distinct diagrams contributing to the leading ($L - V = 1$) and next-to-leading order ($L - V = 2$) of the $\frac{1}{\epsilon}$ expansion of $Z^{(1)}$. Only the ones with shaded backgrounds arise for the unitary class (only parallel encounters and single sense of traversal permitted, as indicated by arrows); they give mutually cancelling contributions in each order, both for the unitary and the orthogonal classes. The numbers n and n' of orbits in respectively $A \cup C$ (full lines) and $B \cup D$ (dashed lines) are written above each column as $n : n'$.

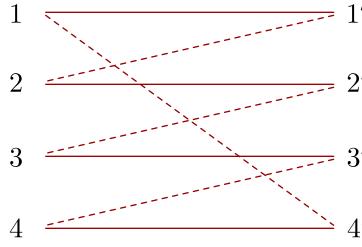


FIG. 4: Topological picture of an encounter where connections $1-1', 2-2', 3-3', 4-4'$ are replaced by $1-4', 2-1', 3-2', 4-3'$.

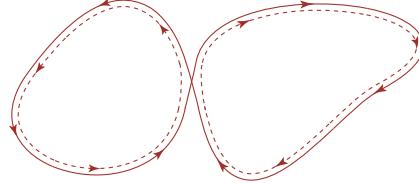


FIG. 5: Simplest (pseudo-)orbits.

C. Classical properties of orbit bunches

We now follow [13, 14] to investigate the classical properties associated to each structure, and then evaluate their contribution to the generating function. We start with the example of orbits differing in a 2-encounter (see Fig. 5). First we need coordinates measuring the phase-space separations between the two encounter stretches. We therefore place a transversal Poincaré surface of section at an arbitrary position inside the encounter and consider the two points where the encounter stretches pierce through that section. The separation of the piercings can be decomposed into a stable component s and an unstable component u [26, 27, 28]. These coordinates determine the difference between the action of the original orbit and the cumulative action of the two partner orbits as $\Delta S = su$ (see also [12, 29]). Moreover, if we define the encounter as the region with phase-space separations $|s|, |u| < c$ (where c is arbitrary but classically small) the time until the end of the encounter can be estimated as $\frac{1}{\lambda} \ln \frac{c}{|u|}$ and the time since the beginning of the encounter is obtained as $\frac{1}{\lambda} \ln \frac{c}{|s|}$. Summation thus yields the overall duration⁶ of the encounter $t_{\text{enc}} \sim \frac{1}{\lambda} \ln \frac{c^2}{|su|}$ [14]. For the semiclassically relevant encounters we need $\Delta S = su \sim \hbar$, i.e., the encounter duration is of the order of the Ehrenfest time $T_E = \frac{1}{\lambda} \ln \frac{c^2}{\hbar}$. Encounters are thus considerably shorter than the orbits, whose periods must be of the order of the Heisenberg time $T_H \propto \hbar^{-1}$.

The likelihood of finding encounters in a periodic orbit can be determined using ergodicity, i.e., the fact that long trajectories and (ensembles of) long periodic orbits uniformly fill the energy shell [30]. Ergodicity implies that the probability for finding two piercings through a Poincaré section with separations t, s, u in differential intervals dt, ds, du is uniform and reads $\frac{dtdsdu}{\Omega}$, with Ω the volume of the energy shell. Statistical independence of the two piercings is assumed here; that assumption is an allowable one since for orbits with a period of the order T_H the links between the encounter stretches typically have durations much longer than the classical timescale needed for the decay of correlations.

Now integration of the probability density $1/\Omega$ yields the number density $w_T(s, u)$ for finding, inside an orbit of period T , 2-encounters with stable and unstable separations s and u . That density was determined in [14, 31] as

$$w_T(s, u) = \frac{T \int dt}{\Omega t_{\text{enc}}(s, u)} . \quad (31)$$

Here the factor T is the period of the orbit (originating from integration over the time of the first piercing). $\int dt$ is an integral over the time of the second piercing reckoned from the first one, with the following restriction: The

⁶ We note that for general hyperbolic systems (where Eq. (30) is only valid in the limit of large times) our formula for t_{enc} is just an approximation. The treatment can be made more rigorous if one avoids this approximation, by following the lines of [25] and Appendix B of [14].

time separation between the two piercings must be at least t_{enc} so that in between the piercings the orbit can go through the rest of the first encounter stretch, a link of positive duration, and the beginning of the second stretch. This reasoning leads to $\int dt = T - 2t_{\text{enc}}$, but we prefer to just write $\int dt$.

Now let us consider more complicated structures with, say, V different encounters inside or between, say, n orbits of periods T_1, T_2, \dots, T_n . The encounters are labelled by an index σ running from 1 to V and the σ th encounter may have $l(\sigma)$ stretches. Drawing a Poincaré section in each encounter we face altogether $L = \sum_{\sigma} l(\sigma)$ piercings, $l(\sigma)$ ones for the σ th encounter. The separations between the piercings belonging to one encounter are characterized by $l(\sigma) - 1$ stable and unstable coordinates $s_{\sigma m}, u_{\sigma m}$ with $m = 1, 2, \dots, l(\sigma) - 1$ (see [14] for the precise definition). These coordinates determine the action difference as $\Delta S = \sum_{\sigma, m} s_{\sigma m} u_{\sigma m}$ and the duration of the σ -th encounter as $t_{\text{enc}, \sigma} = \frac{1}{\lambda} \ln \frac{c^2}{\max_m |s_{\sigma m}| \max_{m'} |u_{\sigma m'}|}$ [14]. By following the same steps as for (31) the number density of encounters is generalized to

$$w_{T_1, T_2, \dots, T_n}(s, u) = \frac{\prod_{i=1}^n T_i \int d^{L-n} t}{\prod_{\sigma} \Omega^{l(\sigma)-1} t_{\text{enc}, \sigma}} \quad (32)$$

with $s = \{s_{\sigma m}\}$, $u = \{u_{\sigma m}\}$. Here the orbit periods T_i arise from integrating over the time of the first piercing in every orbit, and the integral $\int d^{L-n} t$ goes over the possible times of the remaining piercings reckoned from the first one.

D. Contribution of orbit bunches to Z

We can now investigate the impact orbit bunches have on spectral statistics. We thus evaluate the generating function $Z^{(1)}$ as given by the fourfold pseudo-orbit sum in (17), taking into account quadruplets with $A \cup C$ and $B \cup D$ differing in any number of encounters. As before $A \cup C$ and $B \cup D$ may also coincide in any number of component orbits, and therefore $Z_{\text{diag}}^{(1)}$ still arises as a *factor*. We can thus rewrite (17) as

$$Z^{(1)} = Z_{\text{diag}}^{(1)} (1 + Z_{\text{off}}^{(1)}), \quad (33)$$

with $Z_{\text{off}}^{(1)}$ picking up non-empty pseudo-orbits $A \cup C$ and $B \cup D$ that differ in encounters while no longer comprising identically repeated orbits,

$$\begin{aligned} Z_{\text{off}}^{(1)} &= \sum_{\substack{A, B, C, D \\ \text{diff. in enc.}}} \left\langle F_A F_B^* F_C F_D^* (-1)^{n_C + n_D} e^{i\Delta S/\hbar} e^{i(T_A \epsilon_A + T_B \epsilon_B + T_C \epsilon_C + T_D \epsilon_D)/T_H} \right\rangle \\ &\sim \sum_{\substack{A, B, C, D \\ \text{diff. in enc.}}} \left\langle |F_A|^2 |F_C|^2 (-1)^{n_C + n_D} e^{i\Delta S/\hbar} e^{i(T_A \epsilon_A + T_B \epsilon_B + T_C \epsilon_C + T_D \epsilon_D)/T_H} \right\rangle. \end{aligned} \quad (34)$$

The off-diagonal part $Z_{\text{off}}^{(1)}$ of the generating function can now be written as a sum over structures, and then over all quadruplets of pseudo-orbits pertaining to each structure. We thus obtain

$$\begin{aligned} Z_{\text{off}}^{(1)} &= \sum_{\text{structures}} \sum_{\substack{(A, B, C, D) \\ \in \text{structure}}} \frac{(-1)^{n_C + n_D}}{V! \prod_{\sigma} l(\sigma)} \\ &\times \left\langle |F_A|^2 |F_C|^2 e^{i\Delta S/\hbar} e^{i(T_A \epsilon_A + T_B \epsilon_B + T_C \epsilon_C + T_D \epsilon_D)/T_H} \right\rangle, \end{aligned} \quad (35)$$

where as explained above we divided out the number of possible orderings of encounters and stretches, $V! \prod_{\sigma} l(\sigma)$, since otherwise each quadruplet would be counted multiple times. The sum over pseudo-orbits permitted by a structure can be done by summing over all choices for the n component orbits p_1, p_2, \dots, p_n of $A \cup C$ and integrating over the $L - V = \sum_{\sigma} (l(\sigma) - 1)$ pairs of stable and unstable coordinates characterizing the V encounters, with the number density (32) as weight,

$$\begin{aligned} Z_{\text{off}}^{(1)} &= \sum_{\text{structures}} \frac{(-1)^{n_C + n_D}}{V! \prod_{\sigma} l(\sigma)} \sum_{p_1, p_2, \dots, p_n} \prod_{i=1}^n |F_{p_i}|^2 \left\langle \int d^{L-V} s d^{L-V} u \right. \\ &\quad \times \left. w_{T_1, T_2, \dots, T_n}(s, u) e^{i \sum_{\sigma, m} s_{\sigma m} u_{\sigma m} / \hbar} e^{i(T_A \epsilon_A + T_B \epsilon_B + T_C \epsilon_C + T_D \epsilon_D)/T_H} \right\rangle. \end{aligned} \quad (36)$$

Each of the n orbit sums can now be done using the sum rule of Hannay and Ozorio de Almeida, $\sum_{p_i} |F_{p_i}|^2 \rightarrow \int_{T_0}^{\infty} \frac{dT}{T}$. Importing the density w from (32) we see the orbit periods in w cancel against those brought in by the sum rule and face an L -fold time integral, over the n periods and $L - n$ piercing times.

To simplify the foregoing expression for $Z_{\text{off}}^{(1)}$ we change the L -fold time integral just mentioned to an integral over the L link durations. The Jacobian of that transformation is unity. We thus obtain

$$Z_{\text{off}}^{(1)} = \sum_{\text{structures}} \frac{(-1)^{n_C+n_D}}{V! \prod_{\sigma} l(\sigma)} \left\langle \left(\prod_{\sigma=1}^V \int d^{l(\sigma)-1} s_{\sigma} d^{l(\sigma)-1} u_{\sigma} \frac{T_H^{l(\sigma)}}{\Omega^{l(\sigma)-1} t_{\text{enc},\sigma}} e^{i \sum_{m=1}^{l_{\sigma}-1} s_{\sigma m} u_{\sigma m} / \hbar} \right) \times \int_0^{\infty} d^L t_{\text{link}} \frac{1}{T_H^L} e^{i(T_A \epsilon_A + T_B \epsilon_B + T_C \epsilon_C + T_D \epsilon_D) / T_H} \right\rangle, \quad (37)$$

where we inserted compensating powers of the Heisenberg time. The summand coming from a structure in (37) then becomes a product over all links times a product over all encounters. It is easy to evaluate the pertinent factors.

Each link appears precisely once in $A \cup C$ as well as in $B \cup D$. The integral over its duration picks up the integrand $\frac{1}{T_H} e^{i(\epsilon_A \text{ or } C + \epsilon_B \text{ or } D) t_{\text{link}} / T_H}$ and contributes

$$\frac{i}{\epsilon_A \text{ or } C + \epsilon_B \text{ or } D} \quad \text{link factor}; \quad (38)$$

clearly, the index alternatives have to be decided according to which pseudo-orbit the link pertains to before reconnection (A or C) and after reconnection (B or D).

We now turn to the factor coming from an l -encounter. In the original pseudo-orbit $A \cup C$, each of the l stretches may belong to either A or C . We denote by l_A the number of stretches belonging to A and by l_C the number of stretches belonging to C , with $l_A + l_C = l$. When switching connections inside the encounter these stretches are replaced by l new stretches differently connecting the same ports as the old stretches. These new stretches belong to either B or D , and the corresponding numbers are denoted by l_B and l_D with $l_B + l_D = l$. The encounter now yields a contribution l_{enc} to the duration of pseudo-orbit A ; analogously for B, C and D . Hence the encounter draws the part $e^{i(l_A \epsilon_A + l_B \epsilon_B + l_C \epsilon_C + l_D \epsilon_D) t_{\text{enc}} / T_H}$ from the phase factor $e^{i(T_A \epsilon_A + T_B \epsilon_B + T_C \epsilon_C + T_D \epsilon_D) / T_H}$ in (35). Collecting everything from (37) pertaining to the encounter under consideration we get the dimensionless term

$$T_H^l \int \frac{d^{l-1} s d^{l-1} u}{\Omega^{l-1} t_{\text{enc}}(s, u)} e^{i \sum_m s_m u_m / \hbar} e^{i(l_A \epsilon_A + l_B \epsilon_B + l_C \epsilon_C + l_D \epsilon_D) t_{\text{enc}}(s, u) / T_H}. \quad (39)$$

For the calculation of the remaining integral we refer to [14, 32] and here but briefly sketch the argument. The key is to Taylor expand the second exponential and show that only the term linear in t_{enc} survives in the limit $\hbar \rightarrow 0$. In that surviving term the encounter duration t_{enc} cancels and the factor $i(l_A \epsilon_A + l_B \epsilon_B + l_C \epsilon_C + l_D \epsilon_D) / T_H$ can be pulled out from the integral. Using $T_H = \frac{\Omega}{2\pi\hbar}$ and $(\frac{T_H}{\Omega})^{l-1} \int d^{l-1} s d^{l-1} u e^{i \sum_m s_m u_m / \hbar} = \left(\int \frac{ds du}{2\pi} e^{isu} \right)^{l-1} \rightarrow 1$ we get

$$i(l_A \epsilon_A + l_B \epsilon_B + l_C \epsilon_C + l_D \epsilon_D) \quad \text{encounter factor}. \quad (40)$$

The sum over structures (37) thus simplifies to

$$Z_{\text{off}}^{(1)} = \sum_{\text{structures}} \frac{1}{V! \prod_{\sigma} l(\sigma)} (-1)^{n_C+n_D} \frac{\prod_{\text{enc}} i(l_A \epsilon_A + l_B \epsilon_B + l_C \epsilon_C + l_D \epsilon_D)}{\prod_{\text{links}} (-i(\epsilon_A \text{ or } C + \epsilon_B \text{ or } D))}. \quad (41)$$

A further simplification can be achieved after inspecting more closely the encounter contributions. For an l -encounter of some structure we include the overcounting factor $\frac{1}{l}$ and face the contribution $\frac{1}{l} i(l_A \epsilon_A + l_B \epsilon_B + l_C \epsilon_C + l_D \epsilon_D)$. Each of the l structures obtained from the given one by cyclic renumbering of the stretches of the selected encounter contributes identically, and altogether these l structures contribute $i(l_A \epsilon_A + l_B \epsilon_B + l_C \epsilon_C + l_D \epsilon_D)$. We now claim that the same result is obtained if we choose to assign to each structure the encounter factor $i(\epsilon_A \text{ or } C + \epsilon_B \text{ or } D)$ where the index alternatives must be decided according to the placement of the *first entrance port*; e.g., for a structure where the entrance port of the first encounter stretch belongs to the pseudo-orbit A before reconnection and to D afterwards we would write $i(\epsilon_A + \epsilon_D)$. Taking a sum over our group of l equivalent structures we then obtain $\sum \epsilon_A \text{ or } C = l_A \epsilon_A + l_C \epsilon_C$ since the first entrance belongs to A and C in l_A and l_C members of the group, respectively. Similarly we have $\sum \epsilon_B \text{ or } D = l_B \epsilon_B + l_D \epsilon_D$. Consequently, the encounter contribution $\frac{1}{l} i(l_A \epsilon_A + l_B \epsilon_B + l_C \epsilon_C + l_D \epsilon_D)$ can be replaced by $i(\epsilon_A \text{ or } C + \epsilon_B \text{ or } D)$ since after summation over the group of equivalent structures associated with the encounter the two expressions give the same result. Eq. (41) can thus be replaced by the simpler formula in which the numerator refers to the first entrance port of each encounter,

$$Z_{\text{off}}^{(1)} = \sum_{\text{structures}} \frac{1}{V!} (-1)^{n_C+n_D} \frac{\prod_{\text{enc}} i(\epsilon_A \text{ or } C + \epsilon_B \text{ or } D)}{\prod_{\text{links}} (-i(\epsilon_A \text{ or } C + \epsilon_B \text{ or } D))}. \quad (42)$$

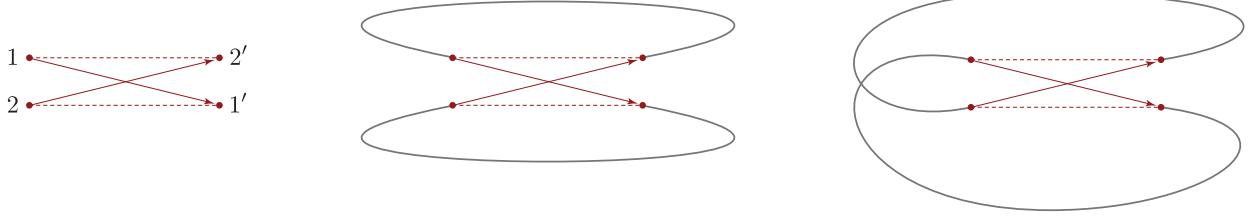


FIG. 6: (a) 2-encounter with ports labelled, exit ports primed; (b, c) two ways of connecting ports with links. Pre-reconnection encounter stretches full lines and post-reconnection ones dashed

A symmetry between link and encounter factors thus results and entails an interesting property of the foregoing “master formula” for $Z_{\text{off}}^{(1)}$: Apart from a factor -1 , each encounter factor is cancelled by the link factor for the link ending at the first entrance port of that encounter. Indeed, the first entrance port and the link ending there belong to the same pseudo-orbit before the reconnection, either A or C , and likewise after the reconnection, either B or D . After that cancellation, the numerator of the ϵ -dependent fraction becomes $(-1)^V$ and only the denominator remains ϵ -dependent. That property will be important for proving the cancellation of all encounter contributions in Sect. V.

According to (42) the contribution of a structure to the generating function is of the order $\frac{1}{\epsilon^{L-V}}$. The sum over structures therefore generates a power series in $\frac{1}{\epsilon}$. Low-order diagrams can easily be drawn; for instance, Fig. 3 depicts all diagrams contributing in the leading two orders, $\frac{1}{\epsilon}$ and $\frac{1}{\epsilon^2}$, both for the unitary and orthogonal class (apart from a further diagram containing two copies of Fig. 5 and a diagram containing one copy of Fig. 5 and one Sieber-Richter pair).

E. Example: Structures for single 2-encounter

It is well to illustrate the formula (42) for the simple example of a single 2-encounter with two entrance and exit ports, depicted in Fig. 6 (a); the full and dashed lines respectively show the encounter stretches before and after reconnection. We can connect the ports by links (stage S2) in $L! = 2$ ways; the two resulting diagrams are shown in parts (b) and (c) of Fig. 6.

We first deal with Fig. 6 (b) where we face a single pre-reconnection orbit γ_0 incorporating all ports and links. After the reconnection the graph disconnects into two orbits γ_1 and γ_2 each of which contains one entrance and one exit port as well as one link; the labels 1 and 2 signal the number of the respective entrance port.

At the stage S3 we generate structures by dividing the orbits among the pseudo-orbits. Since we deal with three orbits there are 2^3 distinct such allotments; four of these are

1. $A = \{\gamma_0\}, B = \{\gamma_1, \gamma_2\}, C = D = \emptyset,$
2. $A = \{\gamma_0\}, D = \{\gamma_1, \gamma_2\}, C = B = \emptyset,$
3. $A = \{\gamma_0\}, B = \{\gamma_1\}, D = \{\gamma_2\}, C = \emptyset,$
4. $A = \{\gamma_0\}, D = \{\gamma_2\}, B = \{\gamma_1\}, C = \emptyset,$

(43)

and four similar structures have A interchanged with C . Consider the first structure in the foregoing list. The entrance port 1 belongs to $\gamma_0 \in A$ before, and to $\gamma_1 \in B$ after reconnection. Therefore the encounter factor in the numerator reads $i(\epsilon_A + \epsilon_B)$. Both links belong to A before and B after reconnection; hence the denominator is $[-i(\epsilon_A + \epsilon_B)]^2$. The factors in (42) are calculated using $V = 1, n_C = n_D = 0$; the structure thus yields $Z_1 = \frac{i}{\epsilon_A + \epsilon_B}$. The contribution of the second structure differs by the replacement $B \rightarrow D$; there are two orbits in D and none in C such that the sign is unchanged and we get $Z_2 = \frac{i}{\epsilon_A + \epsilon_D}$. In the third structure the first entrance port belongs to A before and B after reconnection. The two links are in A before reconnection; after reconnection one of them goes to B , another one to D . The sign factor is negative since $n_C = 0, n_D = 1$. Hence we get $Z_3 = -\frac{i(\epsilon_A + \epsilon_B)}{(\epsilon_A + \epsilon_B)(\epsilon_A + \epsilon_D)} = -\frac{i}{\epsilon_A + \epsilon_D}$. Similarly, the fourth structure gives $Z_4 = -\frac{i(\epsilon_A + \epsilon_D)}{(\epsilon_A + \epsilon_B)(\epsilon_A + \epsilon_D)} = -\frac{i}{\epsilon_A + \epsilon_B}$. Obviously, Z_3 cancels with Z_2 and Z_4 with Z_1 .

The reader is invited to repeat the calculations for the remaining four structures and again find full cancellation. Finally, a patient reader will check that Fig. 6 (c) leads to 8 more structures whose contributions coincide with those of Fig. 6 (b) and also cancel pairwise.

V. CANCELLATION OF OFF-DIAGONAL CONTRIBUTIONS

We now show that for systems without time-reversal invariance all off-diagonal contributions mutually cancel. Our proof is based on two cancellation mechanisms: Each structure is cancelled either by a structure in which the orbits are *partitioned differently into pseudo-orbits* or in which *one link has been removed*.

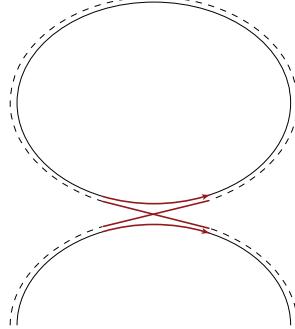


FIG. 7: Case 1 of cancellation: Structure involving orbit with single encounter stretch

Different partition: To start we consider structures where one of orbits (say, in A) only contains *one encounter stretch with two ports and a link*, as in Fig. 7, and *the stretch is the first in its encounter*. The ports and the link then also have to belong to the same partner orbit (included, say, in B). Then according to Eq. (42) the contribution of the structure contains a factor $i(\epsilon_A + \epsilon_B)$ from the entrance port of the stretch and a factor $\frac{1}{-i(\epsilon_A + \epsilon_B)}$ from the link; these link and stretch factors multiply to -1 .

Now we compare with a structure where the orbit from Fig. 7 is included in C rather than A . Then the above factors are replaced by $i(\epsilon_C + \epsilon_B)$ and $\frac{1}{-i(\epsilon_C + \epsilon_B)}$, and the product is again -1 . However, there is an extra factor -1 because the number of orbits in C is increased by one. Hence the two structures considered have cancelling contributions.

The same argument goes through if B is replaced by D , or if the orbit in Fig. 7 is a partner orbit; in the latter case its inclusion in B and D yields mutually cancelling contributions. In all cases we are free to drop from Eq. (42) all structures where an orbit includes just one link and one encounter stretch which is the first in its encounter.

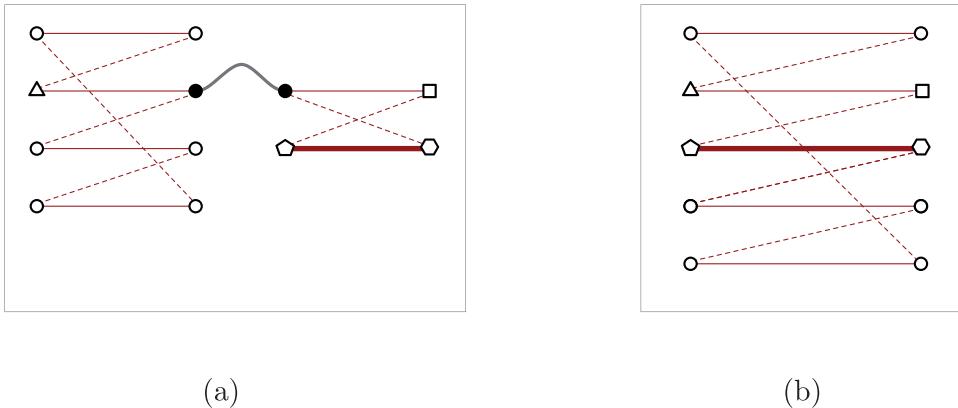


FIG. 8: Case 2 of cancellation: Structure involving 2-encounter and l -encounter cancels against structure with the two encounters merged to an $(l+1)$ -encounter; only the link disappearing in the merger is drawn. Reference stretches in both structures thick

Shrunken links: To get rid of the remaining contributions we consider an arbitrary structure involving a 2-encounter, and there focus specifically on the *link ending at the first entrance port of the 2-encounter*. Then where does that link start? The only possibility that we still have to consider is that it starts at the exit port of a *different* encounter (with an arbitrary number l of stretches). All other possibilities have been dealt with above: if the link would start at the exit port of the same stretch the corresponding original orbit would just include one stretch and one link; if the link would start at the exit of the second stretch, there would be a partner orbit with only one stretch and one link. Hence we must have a situation as depicted schematically in Fig. 8a for $l = 4$.

Now let us consider a structure that looks topologically the same as Fig. 8a but has the link just considered shrunk away. In that structure the two ports connected by the link in question (the black ports in the picture) are ignored, leaving altogether $l + 1$ entrance ports and $l + 1$ exit ports. These ports have to be connected as in Fig. 8a, with connections in the original orbits following the full lines and connections in the partner orbits following the dashed lines. If we leave out the black ports, this leads to connections as in Fig. 8b. We conclude that when the link between the 2-encounter and the l -encounter is removed these two encounters merge into a single $(l + 1)$ -encounter. The contribution of the new structure is proportional to the old one. Just the factor $\frac{1}{i(\epsilon_A \text{ or } C + \epsilon_B \text{ or } D)}$ from the link and the corresponding factor $i(\epsilon_A \text{ or } C + \epsilon_B \text{ or } D)$ from the removed entrance port are gone, and a change of sign is again incurred. Hence the two structures depicted have a chance to cancel.

What remains to be done is a little bit of book-keeping (e.g., to account for the numbering of encounters, and for the possibility that removing links from different structures yields the same result). To do so, we single out one stretch in each structure as a reference; that stretch must not be the first stretch in its encounter. The number of eligible stretches (i.e. the number of all stretches L minus the number of encounters V) will be denoted by n . We now sum over all structures *and* over all possibilities of choosing references stretches inside a structure; to avoid an obvious overcounting we modify the summand of Eq. (42) to include an additional divisor n .

Now for each contribution where the reference stretch belongs to a 2-encounter we can find exactly one contribution with opposite sign where the reference stretch belongs to an l -encounter with $l > 2$. This is done by (i) merging the 2-encounter with a different encounter as described above, (ii) renumbering the encounters to make up for the loss of one encounter (i.e. decrease by 1 the labels of all encounters which are larger than the label of the removed encounter), and (iii) choosing the encounter stretch following the merged one (e.g., the fourth encounter stretch in Fig. 8b) as our new reference stretch. The last step guarantees that the reference stretch ends up in any position but the first in its encounter, as demanded above. It is clear that the new contribution has the same n as the initial one (as both L and V have been reduced by one), hence the additional divisor $1/n$ weighing every contribution remains unchanged.

Each contribution where the reference stretch belongs to an l -encounter can be accessed by following this procedure. This becomes clear if we take a contribution as in Fig. 8b and then undo the steps described above to arrive at Fig. 8a. The same contribution as in Fig. 8b results from V different contributions as in Fig. 8a since the removed 2-encounter may have an arbitrary index between 1 and V . However this degeneracy is compensated by the change of the factorial in the denominator of Eq. (42). Since the remaining factors are the same but the sign is different, we thus see that every contribution where the reference stretch belongs to an l -encounter exactly cancels with V “parent” contributions with the reference stretch in the 2-encounter.

To summarize, we have seen that all off-diagonal contributions cancel due to two mechanisms: For quadruplets where at least one orbit is limited to just one encounter stretch (which is the first in its encounter) and one link, the structures where this orbit is assigned to different pseudo-orbits mutually cancel. For quadruplets where no orbit is of this type we singled out a reference stretch. We then distinguished between contributions where this reference stretch is part of a 2-encounter or an l -encounter with $l > 2$. By the construction of Fig. 8 we were able to show that these two types of contributions mutually cancel. Hence for systems without time-reversal invariance we have $Z_{\text{off}}^{(1)} = 0$ and only the diagonal term remains.

VI. SEMICLASSICAL CONSTRUCTION OF A SIGMA MODEL

An alternative proof for the cancellation of all off-diagonal contributions can be based on field-theoretical techniques. We shall show that the whole series (42) can be summed to become an integral over a certain matrix manifold. That matrix integral turns out to coincide with the one known from the σ -model in RMT, and gives the same result as the diagonal approximation.

A. Matrix elements for ports and contraction lines for links

We recall that structures of pseudo-orbit quadruplets were defined by (i) fixing the numbers of encounters and stretches involved (the stretches being ordered according to the way they are connected in $A \cup C$ and $B \cup D$), (ii) connecting the encounter ports by links, and (iii) distributing the resulting pre-reconnection orbits over the pseudo-orbits A and C , and the post-reconnection orbits between B and D .

Now we encode each structure in (42) by a sequence of alternating symbols B_{kj} and \tilde{B}_{jk} respectively standing for entrance and exit ports. We start with the first entrance port of the first encounter, continue with the first exit port of the first encounter, then list all further ports of that encounter, then of all further encounters. The indices of B, \tilde{B} indicate the affiliation to the pseudo-orbits. Namely, the second index of B and the first index of \tilde{B} determine whether

before reconnection the port belongs to A (case $j = 1$) or to C ($j = 2$). Similarly k shows whether the port belongs after reconnection to B (if $k = 1$) or to D ($k = 2$). Due to the two possible values for both k and j we altogether confront four B_{kj} 's and four \tilde{B}_{jk} 's.

Our conventions for the numbering of ports imply that the ports corresponding to a B and the subsequent \tilde{B} (entrance and exit ports with the same number) are connected by an encounter stretch of an original orbit. Hence they must belong to the same original pseudo-orbit and have the same index j , as in $B_{k'j}\tilde{B}_{jk}$. Similarly the ports associated to a \tilde{B} and the subsequent B (representing an exit port and the entrance port with the following number) must be connected by an encounter stretch of a partner orbit. Consequently they belong to the same partner pseudo-orbit and share the same index k , as in $\tilde{B}_{jk}B_{kj''}$. The same applies for the last \tilde{B} and the first B in each encounter. In any case neighboring indices of successive port symbols must coincide. For a 2-encounter with two entrance and two exit ports as in Fig. 6 (a) we thus get the four-symbol sequence

$$B_{k_1,j_1}\tilde{B}_{j_1,k_2}B_{k_2,j_2}\tilde{B}_{j_2,k_1}, \quad (44)$$

just like in a trace of a matrix product; note that the sub-indices of the indices j, k signal port or encounter stretch labels within the encounter. For a structure with arbitrarily many encounters indexed by $\sigma = 1, 2, \dots, V$ and stretches inside each encounter indexed by $i = 1, 2, \dots, l(\sigma)$ the sequence of ports gives rise to the product

$$\prod_{\sigma=1}^V \left(B_{k_{\sigma 1}, j_{\sigma 1}} \tilde{B}_{j_{\sigma 1}, k_{\sigma 2}} B_{k_{\sigma 2}, j_{\sigma 2}} \dots \tilde{B}_{j_{\sigma l(\sigma)}, k_{\sigma 1}} \right). \quad (45)$$

Note that we do not imply summation over repeated indices either here or anywhere else in the paper. The indices j and k now carry two subindices each, $\sigma = 1, 2, \dots, V$ to label the encounter and $i = 1, 2, \dots, l(\sigma)$ to label the stretches within an encounter. Such “proliferation” of subindices notwithstanding there are still only four different B_{kj} 's and four \tilde{B}_{jk} 's. Whenever possible below we shall suppress the “address labels” on k and j . If we take B_{kj}, \tilde{B}_{jk} to be elements of 2×2 matrices $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$ and similarly for \tilde{B} we may consider the expression above as one of the summands in the expansion of the trace product $\prod_{\sigma=1}^V \text{tr}(B\tilde{B})^{l(\sigma)}$.

The links of a structure will be indicated by drawing “contraction lines” between the respective exit and entrance ports. Each such line connects one entrance and one exit port of some orbit *and* its partner (recall that links do not change when shifting connections in encounters). The indices j, k of B and \tilde{B} indicating accommodation in original and partner pseudo-orbits therefore have to coincide.

Example: In our above example of a product of four symbols the ports can be connected by links in two ways:

$$\boxed{B_{k_1,j_1} \tilde{B}_{j_1,k_2} B_{k_2,j_1} \tilde{B}_{j_1,k_1}} \quad \text{or} \quad \boxed{\tilde{B}_{k_1,j_1} \tilde{B}_{j_1,k_1} B_{k_1,j_2} \tilde{B}_{j_2,k_1}}; \quad (46)$$

note the agreement of indices of the connected ports. We can even complement these diagrams to topological pictures of the orbits. If we draw lower lines to indicate how the ports are connected inside the encounter we obtain

$$\begin{array}{ccc} \boxed{B_{k_1,j_1} \tilde{B}_{j_1,k_2} B_{k_2,j_1} \tilde{B}_{j_1,k_1}} & \text{and} & \boxed{B_{k_1,j_1} \tilde{B}_{j_1,k_2} B_{k_2,j_1} \tilde{B}_{j_1,k_1}} \\ \text{or} \\ \boxed{B_{k_1,j_1} \tilde{B}_{j_1,k_1} B_{k_1,j_2} \tilde{B}_{j_2,k_1}} & \text{and} & \boxed{B_{k_1,j_1} \tilde{B}_{j_1,k_1} B_{k_1,j_2} \tilde{B}_{j_2,k_1}}. \end{array} \quad (47)$$

Here the upper pair represents the structures in the list (43) and pertains to Fig. 6. The first picture shows the single orbit before reconnection (with encounter stretches connecting the ports as $\boxed{B\tilde{B}}$). The second picture depicts the two orbits after reconnection (with encounter stretches connecting the ports as $\boxed{\tilde{B}B}$). Analogously the lower pair represents the structures pertaining to Fig. 6 (c).

Diagrammatic rules: We can now translate the previous diagrammatic rules for the contribution of a structure in (42) to the new language of contracted sequences of matrix elements:

(i) The ports B_{kj} and $\tilde{B}_{j'k'}$ connected by a link must have the same subscripts $j = j'$ and $k = k'$, and the contribution from that link in (42) equals $[-i(\epsilon_A \text{ or } C + \epsilon_B \text{ or } D)]^{-1}$. We can thus write the factor provided by a contraction between $B_{kj}, \tilde{B}_{j'k'}$ as $[-i(\epsilon_j + \epsilon'_k)]^{-1} \delta_{kk'} \delta_{jj'}$ where ϵ_j stands for ϵ_A ($j = 1$) or ϵ_C ($j = 2$); similarly ϵ'_k

equals ϵ_B for $k = 1$ and ϵ_D for $k = 2$. (The primed energy will always refer to the post-reconnection pseudo-orbits B or D .)

(ii) The indices in an encounter factor $i(\epsilon_{A \text{ or } C} + \epsilon_{B \text{ or } D})$ in (42) are determined by the pseudo-orbits containing the first entrance port of the encounter. In (45) the first entrance port of the σ th encounter is denoted by $B_{k_{\sigma 1}, j_{\sigma 1}}$ such that the respective encounter factor can be rewritten as $i(\epsilon_{j_{\sigma 1}} + \epsilon'_{k_{\sigma 1}})$.

(iii) Finally Eq. (42) implies that there should be a factor $(-1)^{n_C + n_D}$ depending on the numbers of orbits included in the pseudo-orbits C and D .

B. Wick's theorem and link summation

We now want to sum over all structures, which in particular involves summation over all possible ways to draw links alias all contractions. A useful tool for that summation is provided by Wick's theorem. Consider the integral

$$\int dbdb^* e^{i\epsilon bb^*} f(b, b^*) \quad (48)$$

over the complex plane with $dbdb^* \equiv \frac{d\text{Re}b d\text{Im}b}{\pi}$ and $f(b, b^*)$ a product involving an equal number of b 's and b^* 's, and $\text{Im}\epsilon > 0$ to ensure convergence. Wick's theorem states (i) that such Gaussian integrals can be written as sums over “diagrams” where each b in $f(b, b^*)$ is connected to one b^* by a “contraction line”, and (ii) that for complex b, b^* these diagrams can be evaluated by performing the Gaussian integral over contracted elements as if the rest of the integrand were absent. The contraction lines can then be eliminated as in

$$\int dbdb^* \prod b b^* g(b, b^*) e^{i\epsilon bb^*} \equiv \frac{1}{-i\epsilon} \int dbdb^* g(b, b^*) e^{i\epsilon bb^*}. \quad (49)$$

Using this contraction rule one can stepwise remove contraction lines and the b 's and b^* 's connected by these lines and thereby in each step earn a factor $-\frac{1}{i\epsilon}$. For complex b and b^* Wick's theorem is just a fancy way of writing down the integral $\int dbdb^* e^{i\epsilon bb^*} (bb^*)^n = \frac{n!}{(-i\epsilon)^{n+1}}$; the factorial gives the number of possible ways to draw contraction lines.

But remarkably, Wick's theorem also holds, up to a sign in Eq. (49), if we replace b and b^* by Grassmann variables η and η^* . Grassmann variables are formally defined as anticommuting, $\eta\eta^* = -\eta^*\eta$, and with integrals $\int d\eta = \int d\eta^* = 0$, $\int d\eta d\eta^* = 1$ [3]. While η and η^* should in general be seen as independent variables, for the purposes of this paper it is convenient to imagine η^* to be the complex conjugate of η and define the complex conjugate of η^* to be $(\eta^*)^* = -\eta$. Imagining $\exp(i\epsilon\eta\eta^*)$ Taylor-expanded we obtain

$$\int d\eta d\eta^* \eta\eta^* e^{i\epsilon\eta\eta^*} \equiv \int d\eta d\eta^* \eta\eta^* e^{i\epsilon\eta\eta^*} = -1 = \frac{1}{i\epsilon} \int d\eta d\eta^* e^{i\epsilon\eta\eta^*}. \quad (50)$$

This is the analogue of (49) with $g = 1$ which is the only case of interest since all powers of the Grassmann variables higher than 1 vanish. Below we shall apply contraction rules of the foregoing types to integrals over an arbitrary number of pairs of commuting and anti-commuting variables. Following traditions of the physical literature we shall refer to the anti-commuting variables as “Fermionic” as opposed to the commuting “Bosonic” ones.

We now want to use Wick's theorem to represent sums over all possible ways of drawing contraction lines between B_{kj} 's and \tilde{B}_{jk} 's, for general products of the form (45). This can be done if we declare these symbols to be either complex Bosonic or Fermionic variables and make the variables B_{kj} and \tilde{B}_{jk} between which we want to draw contraction lines mutually conjugate up to a sign, i.e., we set $\tilde{B}_{jk} \equiv \pm B_{kj}^*$. (The choice of the Bosonic or Fermionic character of the variables and of the sign will be made later.) Then all possible contractions are generated by taking the corresponding expression without contraction lines, and integrating with a Gaussian weight that has a term proportional to $B_{kj}B_{kj}^*$ in the exponent. The prefactor of $B_{kj}B_{kj}^*$ should be chosen as $i(\epsilon_j + \epsilon'_k)$. Then Wick's theorem yields the desired factor $[-i(\epsilon_j + \epsilon'_k)]^{-1}$ for each link belonging to the pseudo-orbits j and k (up to a sign factor arising for Fermionic variables). If we also include the encounter factor $i(\epsilon_{j_{\sigma 1}} + \epsilon'_{k_{\sigma 1}})$ we are led to expressions of the type

$$\pm \int d[B, \tilde{B}] e^{\pm \sum_{j,k=1,2} i(\epsilon_j + \epsilon'_k) B_{kj} B_{kj}^*} \left\{ \prod_{\sigma=1}^V \left(\pm i(\epsilon_{j_{\sigma 1}} + \epsilon'_{k_{\sigma 1}}) B_{k_{\sigma 1}, j_{\sigma 1}} \tilde{B}_{j_{\sigma 1}, k_{\sigma 2}} B_{k_{\sigma 2}, j_{\sigma 2}} \dots \tilde{B}_{j_{\sigma l(\sigma)}, k_{\sigma 1}} \right) \right\} \quad (51)$$

where the integration measure involves the product of all independent differentials,

$$d[B, \tilde{B}] = \frac{1}{\pi^2} d(\text{Re}B_{11}) d(\text{Im}B_{11}) d(\text{Re}B_{22}) d(\text{Im}B_{22}) dB_{12} dB_{12}^* dB_{21} dB_{21}^*. \quad (52)$$

All (convergent) expressions of this type contain all relevant structures with the appropriate link and encounter factors.

C. Signs

The most delicate task is to fix the signs in (51) and the Bosonic vs. Fermionic nature of the variables so as to produce the sign factor $(-1)^{n_C+n_D}$ in the sum over structures (42). We shall make the following choices which are inspired by the supersymmetric sigma model of RMT (see [3] and Appendix B) but will here be justified purely on semiclassical grounds: We take B_{11}, B_{22} Bosonic and B_{12}, B_{21} Fermionic, and let the supermatrices B and \tilde{B} be related as

$$\tilde{B} = \sigma_z B^\dagger = \begin{pmatrix} B_{11}^* & B_{21}^* \\ B_{12}^* & -B_{22}^* \end{pmatrix} \quad (53)$$

where $\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ and the adjoint of a supermatrix is $B^\dagger = \begin{pmatrix} B_{11}^* & B_{21}^* \\ -B_{12}^* & B_{22}^* \end{pmatrix}$. Then we pick the signs in the Gaussian and the product of B 's and \tilde{B} 's of (51) such that these terms can be written in terms of supertraces $\text{Str } X \equiv \text{tr } \sigma_z X$. We also add an overall minus sign.

For the *Gaussian integral* we thus write

$$\begin{aligned} - \int d[B, \tilde{B}] \exp \left(i(\epsilon_1 + \epsilon'_1) B_{11} B_{11}^* - i(\epsilon_1 + \epsilon'_2) B_{21} B_{21}^* + i(\epsilon_2 + \epsilon'_1) B_{12} B_{12}^* + i(\epsilon_2 + \epsilon'_2) B_{22} B_{22}^* \right) \{ \dots \} \\ = - \int d[B, \tilde{B}] \exp \left(i \text{Str } \hat{\epsilon}' B \tilde{B} + i \text{Str } \hat{\epsilon} \tilde{B} B \right) \{ \dots \} \equiv \langle\langle \{ \dots \} \rangle\rangle; \end{aligned} \quad (54)$$

here the curly brackets $\{ \dots \}$ have the same content as in (51); to compact the notation we employ the matrices

$$\hat{\epsilon} = \text{diag}(\epsilon_1, \epsilon_2), \quad \hat{\epsilon}' = \text{diag}(\epsilon'_1, \epsilon'_2); \quad (55)$$

moreover, in the last member of the foregoing chain of equations we abbreviate the B -integral with the Gaussian weight (including the negative sign factor) by double angular brackets $\langle\langle \dots \rangle\rangle$. Since the Gaussian is not normalized, $\langle\langle \dots \rangle\rangle$ is not an average, i.e., $\langle\langle 1 \rangle\rangle \neq 1$. Instead we have

$$\begin{aligned} \langle\langle 1 \rangle\rangle = - \int d[B, \tilde{B}] \exp \left(i(\epsilon_1 + \epsilon'_1) B_{11} B_{11}^* - i(\epsilon_1 + \epsilon'_2) B_{21} B_{21}^* + i(\epsilon_2 + \epsilon'_1) B_{12} B_{12}^* + i(\epsilon_2 + \epsilon'_2) B_{22} B_{22}^* \right) \\ = \frac{(\epsilon_1 + \epsilon'_2)(\epsilon_2 + \epsilon'_1)}{(\epsilon_1 + \epsilon'_1)(\epsilon_2 + \epsilon'_2)} = \frac{(\epsilon_A + \epsilon_D)(\epsilon_C + \epsilon_B)}{(\epsilon_A + \epsilon_B)(\epsilon_C + \epsilon_D)}, \end{aligned} \quad (56)$$

reminiscent of the diagonal part $Z_{\text{diag}}^{(1)}$ of $Z^{(1)}$. The factor $\langle\langle 1 \rangle\rangle$ remains after all contractions are performed and thus all B , \tilde{B} removed. Reassuringly it accounts, once multiplied with $e^{i(\epsilon_A + \epsilon_B - \epsilon_C - \epsilon_D)/2}$, for all diagonal quadruplets.

When writing the *product of B 's and \tilde{B} 's* in (51) in terms of a supertrace, we still need to keep the indices fixed (note that no summation over port labels is implied!). We thus employ projection matrices $P_1 = \text{diag}(1, 0)$, $P_2 = \text{diag}(0, 1)$ and write (51) as

$$\left\langle\left\langle \prod_{\sigma=1}^V i(\epsilon_{j_{\sigma 1}} + \epsilon'_{k_{\sigma 1}}) \text{Str } P_{k_{\sigma 1}} B P_{j_{\sigma 1}} \tilde{B} P_{k_{\sigma 2}} \dots P_{j_{\sigma l(\sigma)}} \tilde{B} \right\rangle\right\rangle. \quad (57)$$

Note that just like for a trace, it is possible to cyclically permute the supermatrices appearing in this supertrace.

If we also install the weight $\frac{1}{V!}$ of a structure and the Weyl factor, we obtain

$$\begin{aligned} Z^{(1)}(V, \{l(\sigma), j_{\sigma i}, k_{\sigma i}\}) &= \frac{e^{i(\epsilon_A + \epsilon_B - \epsilon_C - \epsilon_D)/2}}{V!} \\ &\times \left\langle\left\langle \prod_{\sigma=1}^V i(\epsilon_{j_{\sigma 1}} + \epsilon'_{k_{\sigma 1}}) \text{Str } P_{k_{\sigma 1}} B P_{j_{\sigma 1}} \tilde{B} P_{k_{\sigma 2}} \dots P_{j_{\sigma l(\sigma)}} \tilde{B} \right\rangle\right\rangle \end{aligned} \quad (58)$$

which gives the cumulative contribution to $Z^{(1)} = Z_{\text{diag}}^{(1)}(1 + Z_{\text{off}}^{(1)})$ of all structures with V encounters, each having a fixed number of stretches $l(\sigma), \sigma = 1 \dots V$ and a fixed distribution of ports $\{j_{\sigma i}, k_{\sigma i}\}$ among the pseudo-orbits; each allowed way of drawing contraction lines makes for one structure. The purely diagonal additive term is due to $V = 0$.

The only thing still to be shown for a proof of this statement is that our conventions incorporate the correct sign $(-1)^{n_C+n_D}$ for each structure depending on the number of orbits included in C and D . To fill this gap we reexpress

the contraction rules (49) and (50) in terms of B and \tilde{B} . Let us consider one way of drawing contraction lines in Eq. (58), and moreover select a particular contraction, either between ports within the same encounters (inside the same supertrace) or from different encounters. Then we can prove two identities connecting the contribution of a structure with that of another structure containing two ports and one link less (see A for details of the derivation),

$$\left\langle\!\left\langle \text{Str}(P_k \overbrace{BP_j Y} \text{Str}(XP_{j'} \overbrace{\tilde{B}P_{k'}} \dots) \right\rangle\!\right\rangle = -\frac{\delta_{jj'}\delta_{kk'}}{i(\epsilon_j + \epsilon'_k)} \langle\langle \text{Str}(P_k X P_j Y) \dots \rangle\rangle, \quad (59)$$

$$\left\langle\!\left\langle \text{Str}(P_k \overbrace{BP_j U P_{j'} \tilde{B}P_{k'}} V) \dots \right\rangle\!\right\rangle = -\frac{\delta_{jj'}\delta_{kk'}}{i(\epsilon_j + \epsilon'_k)} \langle\langle \text{Str}(P_j U) \text{Str}(P_k V) \dots \rangle\rangle. \quad (60)$$

In these formulae, contraction lines other than the selected one are not drawn and are assumed to coincide in the left and right hand sides. In terms of periodic orbits, the rules (59, 60) respectively account for the two ports connected by a link belonging to either different encounters or to the same one. The symbols X, Y, U, V represent matrix products as in (58) with $Y = \tilde{B} \dots \tilde{B}$, $X = B \dots B$; on the other hand, $U = \tilde{B} \dots B$, $V = B \dots \tilde{B}$. The terms $-\delta_{jj'}\delta_{kk'} \frac{1}{i(\epsilon_j + \epsilon'_k)}$ are consistent with the diagrammatic rules stated above: Links can only connect ports associated to the same original pseudo-orbit $j = j'$ and the same partner pseudo-orbit $k = k'$, and contribute factors $-\frac{1}{i(\epsilon_j + \epsilon'_k)}$. The contraction rules can be used to stepwise simplify our expressions for the structure contributions. Each step removes two matrices or, in other words, a link connecting two ports from a periodic-orbit structure, without changing other links.

At first sight, the rules (59, 60) do not seem to yield any sign factors. But we must pay special attention to steps where the contraction line pertaining to the *last* link of an orbit is removed, after removal of all other contraction lines of some “parent” orbit in previous steps. The B and \tilde{B} so connected represent the only ports of the remnant orbit. That orbit is still periodic, and therefore the two ports are connected not only by the link but also by an encounter stretch. The two matrices B ’s and \tilde{B} ’s then have just one projection matrix in between. If the orbit in question is an original one, the matrices follow each other like $BP_j \tilde{B}$. We thus have to use contraction rule (60) with $U = 1$. Then the first supertrace on the r.h.s. turns into $\text{Str} P_j$ which equals 1 if $j = 1$ (the orbit belongs to pseudo-orbit A) and -1 if $j = 2$ (the orbit belongs to C). As desired a sign factor -1 thus arises for each orbit included in C .

An analogous result holds for partner orbits. Imagine that we remove the last contraction line associated to a partner orbit. Then the ports connected by this line must also be connected by a partner encounter stretch. The rule (60) then applies with $V = 1$. The second supertrace on the r.h.s. thus equals $\text{Str} P_k$ and yields -1 if $k = 2$, i.e., if the orbit belongs to D .

To conclude, by turning B and \tilde{B} into supermatrices, we have successfully incorporated the sign factor $(-1)^{n_C + n_D}$. Other ways to deal with that sign factor (the replica trick) are discussed in Appendices C and D.

D. Towards a sigma model

To determine $Z^{(1)}$ we sum over all structures of pseudo-orbit quadruplets. In Eq. (58) we had already collected all ways of connecting ports with links. Now we sum over all possibilities of assigning ports to pseudo-orbits, i. e., over the indices $j_{\sigma i}, k_{\sigma i}$ in (58). This means that we have to drop the projection matrices. To keep the factors $i(\epsilon_{\sigma 1} + \epsilon'_{\sigma 1})$ connected to the indices in the supertrace we use the diagonal offset matrices $\hat{\epsilon}, \hat{\epsilon}'$ and write

$$Z^{(1)}(V, \{l(\sigma)\}) = \frac{e^{i(\epsilon_A + \epsilon_B - \epsilon_C - \epsilon_D)/2}}{V!} \left\langle\!\left\langle \prod_{\sigma=1}^V i \text{Str} \left(\hat{\epsilon}(\tilde{B}B)^{l(\sigma)} + \hat{\epsilon}'(B\tilde{B})^{l(\sigma)} \right) \right\rangle\!\right\rangle. \quad (61)$$

Here the term $\text{Str} \hat{\epsilon}'(B\tilde{B})^{l(\sigma)}$ contains the factor $\epsilon'_{k_{\sigma 1}}$ from (58) while $\text{Str} \hat{\epsilon}(\tilde{B}B)^{l(\sigma)}$ includes $\epsilon_{j_{\sigma 1}}$. It remains to sum over the number V of encounters and over their sizes $l(\sigma)$, $\sigma = 1, 2, \dots, V$. We let V run from 0 to ∞ , the term $V = 0$ taking care of the purely diagonal contribution, and obtain

$$\begin{aligned} Z^{(1)} &= e^{i(\epsilon_A + \epsilon_B - \epsilon_C - \epsilon_D)/2} \sum_{V=0}^{\infty} \frac{1}{V!} \left\langle\!\left\langle \left[i \text{Str} \sum_{l=2}^{\infty} \left(\hat{\epsilon}(\tilde{B}B)^l + \hat{\epsilon}'(B\tilde{B})^l \right) \right]^V \right\rangle\!\right\rangle \\ &= e^{i(\epsilon_A + \epsilon_B - \epsilon_C - \epsilon_D)/2} \left\langle\!\left\langle \exp \left[i \text{Str} \sum_{l=2}^{\infty} \left(\hat{\epsilon}(\tilde{B}B)^l + \hat{\epsilon}'(B\tilde{B})^l \right) \right] \right\rangle\!\right\rangle. \end{aligned} \quad (62)$$

Now recall that (62) was derived using a semiclassical approximation, which converges only with large imaginary parts η for all energy offsets. In the limit $\eta \gg 1$ the B -integral draws overwhelmingly dominant contributions from near $B = 0$ whereas other B only make for exponentially small corrections of order $e^{-\eta}$. Therefore in the important

integration region we have $\|B\tilde{B}\| \ll 1$ and the geometric series in the foregoing exponent can be replaced by its sum. Upon explicitly writing the Gaussian integral we get the compact result

$$\begin{aligned} Z^{(1)}(\epsilon_A, \epsilon_B, \epsilon_C, \epsilon_D) &\sim -e^{i(\epsilon_A + \epsilon_B - \epsilon_C - \epsilon_D)/2} \\ &\times \int_{\mathcal{L}} d[B, \tilde{B}] \exp \left[i \text{Str} \left(\hat{\epsilon} \frac{\tilde{B}B}{1 - \tilde{B}B} + \hat{\epsilon}' \frac{B\tilde{B}}{1 - B\tilde{B}} \right) \right] \\ &= - \int_{\mathcal{L}} d[B, \tilde{B}] \exp \left[\frac{i}{2} \text{Str} \left(\hat{\epsilon} \frac{1 + \tilde{B}B}{1 - \tilde{B}B} + \hat{\epsilon}' \frac{1 + B\tilde{B}}{1 - B\tilde{B}} \right) \right], \quad \eta \gg 1. \end{aligned} \quad (63)$$

The integration domain \mathcal{L} must include the vicinity of the point $B = \tilde{B} = 0$ and obey $1 - B\tilde{B} > 0$, to exclude regions where the integrand is exponentially large; otherwise it is arbitrary. Under these limitations and $\eta \gg 1$ the asymptotic $1/\epsilon$ -expansion of the integral (63) coinciding with the semiclassical series (61) is created by the saddle point of the exponent at $B = \tilde{B} = 0$.

If we evaluate (63) in the limit $\eta \gg 1$ (see Appendix B) we obtain the same result $Z^{(1)}$ as in the diagonal approximation, Eq (28). Then the same $Z^{(2)}$ as in the diagonal approximation is obtained using the Riemann-Siegel relation $Z^{(2)}(\epsilon_A, \epsilon_B, \epsilon_C, \epsilon_D) = Z^{(1)}(\epsilon_A, \epsilon_B, -\epsilon_D, -\epsilon_C)$. We see once again that all off-diagonal terms in the generating function cancel.

E. Comparison with the RMT sigma model

In the so called rational parametrization of the zero-dimensional sigma model of RMT (see, e.g., Appendix B) the *exact* generating function Z is given by the integral

$$Z(\epsilon_A, \epsilon_B, \epsilon_C, \epsilon_D) = - \int_{\mathcal{L}_0} d[B, \tilde{B}] \exp \left[\frac{i}{2} \text{Str} \left(\hat{\epsilon} \frac{1 + \tilde{B}B}{1 - \tilde{B}B} + \hat{\epsilon}' \frac{1 + B\tilde{B}}{1 - B\tilde{B}} \right) \right], \quad (64)$$

identical in appearance with (63) and with the integration measure and the structure of the supermatrices the same as there. However there are two important differences: First, the offset variables $\epsilon_{A,B,C,D}$ are all allowed to be real. Second, the complex Bosonic integration variables are assigned specific integration domains, the open unit disk $|B_{11}| < 1$ and the full complex plane $|B_{22}| < \infty$. It is easy to show that one of the eigenvalues of $B\tilde{B}$ is then between 0 and 1 while the other one can take any negative value. Consequently $1 - B\tilde{B} > 0$ and the integration domain \mathcal{L}_0 is in fact a special case of \mathcal{L} in (63).

Now suppose that instead of calculating the integral exactly we consider its high-energy asymptotics. We then get a sum of contributions of the two existing stationary points of the exponent. The first stationary point is at $B = 0$; its contribution coincides with our semiclassical result $Z^{(1)}$.

The second stationary “point” associated with the so called Andreev-Altshuler saddle corresponds to $B_{11} = 0, B_{22} \rightarrow \infty$ (see Appendix B). For $\eta \gg 1$ its contribution is exponentially small compared to the first one and can be dropped, but it is of the same order for small or zero η . Now importantly this contribution is related to $Z^{(1)}$ by $Z^{(2)}(\epsilon_A, \epsilon_B, \epsilon_C, \epsilon_D) = Z^{(1)}(\epsilon_A, \epsilon_B, -\epsilon_D, -\epsilon_C)$ (see Appendix B). This is the same relation as between the two summands in our semiclassical approximation of the generating function, based on the Riemann-Siegel lookalike formula and thus unitarity. As anticipated the generating function and the two-point correlation function of fully chaotic dynamics without time-reversal invariance thus agree with those derived from the Gaussian Unitary Ensemble of RMT.

The following extension of the comparison with the RMT sigma model requires some more knowledge of that model. Previously uninitiated readers may want to first consult with introductions provided in Appendix B and in [3] or with the pioneering work of *K. Efetov* [4].

Q-integral and geometrical interpretation: We first compact the RMT matrix integral (64) to

$$Z = \int dQ e^{\frac{1}{2} \text{Str} Q \left(\begin{smallmatrix} \hat{\epsilon}' & 0 \\ 0 & -\hat{\epsilon} \end{smallmatrix} \right)}; \quad (65)$$

here the 4×4 supermatrix $Q = iT\Lambda^{(1)}T^{-1}$ is related to a diagonal matrix (called “standard saddle”) $\Lambda^{(1)} = \text{diag}(1, 1, -1, -1)$ by conjugation with the supermatrix $T = \begin{pmatrix} 1 & B \\ \tilde{B} & 1 \end{pmatrix}$. The matrices Q and $\Lambda^{(1)}$ obey the nonlinear constraints $-Q^2 = (\Lambda^{(1)})^2 = 1$. The integration domain for the Bosonic elements in Q can be shown to be the two-sphere S_2 and (one of the disjoint halves of) the two dimensional hyperboloid H_2 . In particular, the complex

variable B_{22} in the foregoing rational parametrization can be identified as a stereographic projection variable for the sphere S_2 . The point $B_{22} = 0$ corresponds to the “south pole” while the antipodal “north pole” is approached as $B_{22} \rightarrow \infty$. The complex variable B_{11} , on the other hand, is associated with H_2 .

The foregoing rational parametrization of the manifold of matrices Q is just one of many. In fact, the manifold in question can be characterized without specifying any “coordinates”, and we would like to just state that characterization here, referring to Appendix B of the present paper for some details. For a geometrical interpretation of the matrix integral it is well to reorder the lines and columns in the 4×4 matrices Q, T, Λ as $1234 \rightarrow 1324$, so as to be able to employ 2×2 blocks as $Q = \begin{pmatrix} Q_{BB} & Q_{BF} \\ Q_{FB} & Q_{FF} \end{pmatrix}$; the Bose-Bose block Q_{BB} and the Fermi-Fermi Block Q_{FF} contain only commuting entries while all anticommuting entries are assembled in the off-diagonal blocks Q_{BF}, Q_{FB} . In that “Bose-Fermi notation” (the previously used ordering is known as the advanced-retarded notation) [3] the integration domain for the blocks Q_{BB}, Q_{FF} in the sigma model integral become

$$Q_{BB} \in H_2 \equiv \mathrm{U}(1,1)/\mathrm{U}(1) \times \mathrm{U}(1), \quad Q_{FF} \in S_2 \equiv \mathrm{U}(2)/\mathrm{U}(1) \times \mathrm{U}(1), \quad (66)$$

where $\mathrm{U}(2)$ is the 2-dimensional unitary group, and $\mathrm{U}(1,1)$ is the group of 2-dimensional pseudo-unitary matrices, i.e. the group of matrices obeying $U^\dagger \sigma_z U = \sigma_z$, where $\sigma_z = \mathrm{diag}(1, -1)$. Geometrically, H_2 and S_2 define the two-dimensional hyperboloid and the two-sphere mentioned before.

Andreev-Altshuler saddle point [33, 34, 35]: We now comment on another specific parametrization of the RMT sigma model connected to the contributions $Z^{(2)}$ in our semiclassical approach. This parametrization is obtained from the above rational one by the transformation $Q = RQ'R$ with R a permutation matrix interchanging rows and columns as $1234 \rightarrow 1432$ and obeying $R^2 = 1$. The pertinent Jacobian is unity and the integration manifold is unchanged. The exact generating function can thus be written as the “ Q' -integral”

$$Z = \int dQ' e^{\frac{1}{2} \mathrm{Str} Q' R \begin{pmatrix} \hat{\epsilon}' & 0 \\ 0 & -\hat{\epsilon} \end{pmatrix} R} \quad (67)$$

and the rational parametrization chosen for T in $Q' = T\Lambda^{(1)}T^{-1}$. We have thus recovered the Weyl symmetry, $Z(\epsilon_A, \epsilon_B, \epsilon_C, \epsilon_D) = Z(\epsilon_A, \epsilon_B, -\epsilon_D, -\epsilon_C)$, since conjugation of the matrix of energy offsets with R just amounts to $\epsilon_C \leftrightarrow -\epsilon_D$. On the other hand, it is most illuminating to check that the “point” $B = \tilde{B} = 0$ for Q' corresponds to $Q = R\Lambda^{(1)}R \equiv \Lambda^{(2)}$, the so called Andreev-Altshuler saddle. The two diagonal “saddles” are not within reach of one another through the rational parametrization with any finite B, \tilde{B} . Rather, $\Lambda^{(2)}$ can only be approached as $T\Lambda^{(1)}T^{-1}$ for $B, \tilde{B} \rightarrow \infty$ and vice versa. In particular, south and north pole of the sphere S_2 are swapped in the Q' -integral relative to the Q -integral. In that sense one may say that the Q -integral (65) rationally parametrizes Q w.r.t. the standard saddle while the Q' -integral (67) rationally parametrizes Q w.r.t. the Andreev-Altshuler saddle. We would like to emphasize that either variant of the rational parametrization yields, upon exact evaluation of the superintegral, the full generating function; the inaccessibility of the antipodal saddle does not matter since all of its neighborhood is accounted for.

As already stated, our semiclassical construction of the $[B, \tilde{B}]$ -integral corresponds to choosing the rational parametrization according to the Q -integral in (65), since that choice for the RMT sigma model integral does yield $Z^{(1)}$ in a Gaussian approximation around $B = \tilde{B} = 0$. Had we done our semiclassical analysis for $Z^{(2)}(\epsilon_A, \epsilon_B, \epsilon_C, \epsilon_D)$, we would have ended up with a $[B, \tilde{B}]$ -integral corresponding, in the RMT context, to the alternative rational parametrization in the Q' -integral in (67): The Gaussian approximation about $B = \tilde{B} = 0$ in (67) does yield $Z^{(2)}(\epsilon_A, \epsilon_B, \epsilon_C, \epsilon_D)$, as is easily checked. In other words, the asymptotic contribution which stems from the infinitely remote integration point in the parametrization using the standard saddle, becomes the contribution of the stationary point $B = \tilde{B} = 0$ in the Andreev-Altshuler parametrization and vice versa.

VII. TIME-REVERSAL INVARIANT SYSTEMS

We now propose to generalize the above to time-reversal invariant systems. For them, the possibility to revert orbits (and their pieces) in time makes room for more constructive interference. We shall see that the contributions of orbits differing in encounters no longer cancel but rather generate terms of arbitrarily high orders in $1/\epsilon$. The summation of all relevant pseudo-orbit quadruplets can again be done with the help of a sigma-model type matrix integral.

A. Diagonal approximation and off-diagonal corrections

For time-reversal invariant systems the orbits of (A, C) can not only be repeated in (B, D) but may also be reversed in time and only then included in (B, D) . Thus if we again formulate the *diagonal approximation* in terms of an

exponentiated double sum over orbits (26), the relevant contributions to the exponent involve pairs of orbits which are either identical or mutually time-reversed. This doubling of orbit pairs leads to an additional factor 2 in the exponent, and the resulting term $\frac{(\epsilon_A + \epsilon_D)(\epsilon_C + \epsilon_B)}{(\epsilon_A + \epsilon_B)(\epsilon_C + \epsilon_D)}$ in Eq. (28) needs to be squared.

Moreover, we now have to consider *encounters* where some of the stretches are almost mutually time-reversed rather than close in phase space. For instance, Sieber-Richter pairs as depicted in Fig. 2 contain one encounter with two almost time-reversed stretches. If we change connections inside this encounter, the two links outside the encounter still look almost the same in configuration space, but one of them must be traversed with opposite sense of motion, and this is possible only if the dynamics is time reversal invariant.

Since time reversal changes the sides where a stretch enters or leaves an encounter, it becomes cumbersome to work with the notion of entrance and exit ports. Instead we arbitrarily refer to the ports on one side of each encounter as “left” ports and those on the opposite side as “right” ports. Encounter stretches may lead either from left to right or from right to left. The links may connect left and left, right and right, or left and right ports.

Our formal definition of structures is then changed as follows: S1: When numbering ports we now make sure that in $A \cup C$ the i -th left port is connected to the i -th right port, whereas in $B \cup D$ it is connected to the $(i-1)$ -st right port. S2: When connecting ports through links we have to allow for arbitrary connections between ports on either side of the encounters. S3: Apart from dividing the orbits before reconnection between the pseudo-orbits A and C and the orbits after reconnection among B and D we have to choose the sense of traversal of every orbit.

Again, several structures can describe the same quadruplet. In particular, we decide arbitrarily which side of each encounter is chosen as left or right. For V encounters there are 2^V equivalent possibilities of labelling the sides. To avoid overcounting, this factor needs to be divided out.

The rest of our argument proceeds in analogy to Section IV. Our semiclassical *generating function* turns into

$$\begin{aligned} Z_{\text{diag}}^{(1)} &= e^{i(\epsilon_A + \epsilon_B - \epsilon_C - \epsilon_D)/2} \left(\frac{(\epsilon_A + \epsilon_D)(\epsilon_C + \epsilon_B)}{(\epsilon_A + \epsilon_B)(\epsilon_C + \epsilon_D)} \right)^2 \\ Z_{\text{off}}^{(1)} &= \sum_{\text{structures}} \frac{1}{2^V V!} (-1)^{n_C + n_D} \frac{\prod_{\text{enc}} i(\epsilon_A \text{ or } C + \epsilon_B \text{ or } D)}{\prod_{\text{links}} (-i(\epsilon_A \text{ or } C + \epsilon_B \text{ or } D))} \end{aligned} \quad (68)$$

where the subscripts in the encounter factor depend on the first left port.

We note that some of these structures already appear for systems without time-reversal invariance, or coincide with such structures up to the directions of motion in some of the orbits. One can show that the proof of section V generalizes accordingly, i.e., all structures trivially related to those in the unitary case mutually cancel.

Let us use the semiclassical expansion (68) to calculate the *leading-order oscillatory contribution* to the correlation function. The oscillatory term is obtained by application of $\partial^2/\partial\epsilon_A\partial\epsilon_B$ to

$$\begin{aligned} Z^{(2)}(\epsilon_A, \epsilon_B, \epsilon_C, \epsilon_D) &= Z^{(1)}(\epsilon_A, \epsilon_B, -\epsilon_D, -\epsilon_C) \\ &= e^{\frac{i}{2}(\epsilon_A + \epsilon_B + \epsilon_C + \epsilon_D)} \left[\frac{(\epsilon_A - \epsilon_C)(\epsilon_B - \epsilon_D)}{(\epsilon_A + \epsilon_B)(\epsilon_C + \epsilon_D)} \right]^2 \left(1 + Z_{\text{off}}^{(1)}(\epsilon_A, \epsilon_B, -\epsilon_D, -\epsilon_C) \right) \end{aligned} \quad (69)$$

and then taking the limit $\epsilon_A, \epsilon_B, \epsilon_C, \epsilon_D \rightarrow \epsilon$. The factor $[(\epsilon_A - \epsilon_C)(\epsilon_B - \epsilon_D)]^2$ tends to zero in this limit even after taking the derivatives. Therefore the oscillatory contribution to the correlator is purely off-diagonal. It must be connected to summands in $Z_{\text{off}}^{(1)}(\epsilon_A, \epsilon_B, -\epsilon_D, -\epsilon_C)$ which tend to infinity like $[(\epsilon_A - \epsilon_C)(\epsilon_B - \epsilon_D)]^{-1}$, since multiplication with such terms and differentiation w.r.t. ϵ_A and ϵ_B is needed to eliminate the factor $[(\epsilon_A - \epsilon_C)(\epsilon_B - \epsilon_D)]^2$. The respective contributions in $Z_{\text{off}}^{(1)}(\epsilon_A, \epsilon_B, \epsilon_C, \epsilon_D)$ behave like $[(\epsilon_A + \epsilon_D)(\epsilon_C + \epsilon_B)]^{-1}$. Since inverse powers of the energy offsets can only originate from the link factors, such terms can only be due to quadruplets that contain a link belonging to the pseudo-orbit A before and to D after reconnection, and another link belonging to C before and to B after the reconnection. In particular this means that the corresponding quadruplets must contain at least two orbits, both before and after reconnection.

In the lowest order these are the quadruplets depicted in column 2:2 of Fig. 3. (We omitted quadruplets containing two copies of Fig. 5 or one copy of Fig. 5 and one Sieber-Richter pair. These quadruplets are irrelevant since the contribution of a disconnected diagram can be written as the product of its components, and the factor due to Fig. 5 vanishes by the argument in section IV E.) Among them the only diagram that explicitly requires time-reversal invariance is the one consisting of two Sieber-Richter pairs (the middle diagram in column 2:2 of Fig. 3). In order to satisfy the above conditions on membership in pseudo-orbits we only need to consider the situation where one of these Sieber-Richter pairs involves one orbit included in A and one orbit included in D , whereas the other pair involves one orbit included in B and one orbit included in C . There are $2^4 = 16$ ways to choose directions of motion in each of the four orbits involved, and two choices for considering either of the two encounters as the first, leaving altogether

32 equivalent structures. Their total contribution to $Z_{\text{off}}^{(1)}(\epsilon_A, \epsilon_B, \epsilon_C, \epsilon_D)$ is

$$32 \times \frac{1}{2^2 2!} \frac{[i(\epsilon_A + \epsilon_D)][i(\epsilon_C + \epsilon_B)]}{[-i(\epsilon_A + \epsilon_D)]^2 [-i(\epsilon_C + \epsilon_B)]^2} = -\frac{4}{(\epsilon_A + \epsilon_D)(\epsilon_C + \epsilon_B)};$$

The leading oscillatory contribution to the correlator thus reads

$$C_{\text{double-SR}}(\epsilon) = -2 \frac{\partial^2}{\partial \epsilon_A \partial \epsilon_B} \left[\left(-\frac{4}{(\epsilon_A - \epsilon_C)(\epsilon_B - \epsilon_D)} \right) \times e^{\frac{i}{2}(\epsilon_A + \epsilon_B + \epsilon_C + \epsilon_D)} \frac{(\epsilon_A - \epsilon_C)^2 (\epsilon_B - \epsilon_D)^2}{(\epsilon_A + \epsilon_B)^2 (\epsilon_C + \epsilon_D)^2} \right]_{\parallel} = \frac{1}{2\epsilon^4} e^{i2\epsilon} \quad (70)$$

which coincides with the predictions of RMT.

B. Semiclassical construction of a sigma model

To go beyond the leading order, we again write our semiclassical results in terms of a matrix integral. The derivation of this matrix integral is very similar to the unitary case. First we encode every structure by a sequence of symbols, each symbol standing for a port, and links depicted by contraction lines. Second, we replace every symbol by a Bosonic or Fermionic variable and integrate the product with a suitably chosen Gaussian weight. The contribution of every structure will then coincide with the corresponding contractions in the resulting Gaussian integral; the sum over all possible link connections will be provided by the integral itself. Finally, summing over all encounter sets we get an integral representation of the generating function known from the zero dimensional sigma model of RMT.

1. Matrix elements for ports and contraction lines for links

We again write down the sequence of ports starting with the first left port of the first encounter, continuing with the first right port, etc. The left ports will be denoted by $B_{\nu k, \mu j}$ and the right ports by $\tilde{B}_{\mu j, \nu k}$. The first and the second pair of indices of B respectively refer to the properties of the port before and after reconnection, and vice versa for \tilde{B} . As in the unitary case the Latin index $j = 1, 2$ indicates whether the port belongs to the original pseudo-orbit A ($j = 1$), or C ($j = 2$). The index k reveals whether after reconnection the port belongs to the partner pseudo-orbit B ($k = 1$) or D ($k = 2$).

The Greek indices μ, ν account for the directions of motion through the port and the attached encounter stretches, both in the pertinent original and partner orbits: If $\mu = 1$ the direction is from left to right in the original orbit; if $\mu = 2$ it is from right to left. The index $\nu = 1, 2$ indicates the same for the partner orbits. The symbols $B_{\nu k, \mu j}, \tilde{B}_{\mu j, \nu k}$ can be considered as elements of 4×4 matrices B, \tilde{B} ; to simplify the notation we shall often substitute a capital Latin letter for the composite index like $J = (\mu, j)$.

As for systems without time-reversal invariance the ports are ordered such that an element of B and the immediately following \tilde{B} represent ports that are connected by an encounter stretch of an original orbit. This means that they have to belong to the same original pseudo-orbit, and that the direction of motion (left to right or right to left) must be the same as well. Hence their indices $J = (\mu, j)$ have to coincide. Similarly each \tilde{B} and the immediately following B represent ports that are connected by a stretch of a partner orbit, and the corresponding subscripts $K = (\nu, k)$ coincide. (When a \tilde{B} represents the last right port in its encounter, its subscripts ν, k have to coincide with those of the first B .) Thus the arrangement of indices J and K is the same as in a product of matrices and we obtain a product of the same form as in Eq. (45) but with j and k replaced by J, K .

We now build in the links and depict them by ‘‘contraction lines’’ above the symbol sequences. Since for time-reversal invariant systems links can be drawn between ports on either side of the encounters, the contraction lines can now connect not only B to \tilde{B} but also B to B and \tilde{B} to \tilde{B} . Any ports connected by a link must belong to the same pseudo-orbit, before and after reconnection and hence their Latin indices must coincide, as in the unitary case. The Greek direction indices require new reasoning. A contraction between a \tilde{B} and a B stands for a link connecting a right port to a left port. In this case, if an orbit *leaves* one encounter at the right port it must *enter* the other encounter stretch at the left port. The direction of motion (left to right) is thus the same for both encounter stretches. The same applies for a direction of motion from right to left, and for original and partner orbits alike. Hence for contractions between \tilde{B} and B the Greek subscripts indicating the directions of motion must coincide, $\mu_1 = \mu_2, \nu_1 = \nu_2$. On the other hand, in the contractions $B \dots \tilde{B} \tilde{B} \dots \tilde{B}$ the connected ports lie on the same side, and hence their directions

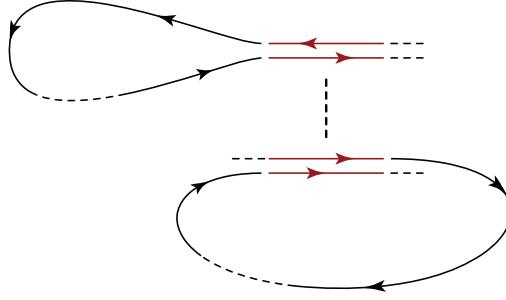


FIG. 9: In the presence of time reversal invariance a link can connect ports on the same side as well as on opposite sides of an encounter. The linked encounter stretches are traversed in antiparallel or parallel direction, correspondingly.

of motion must be opposite. These possibilities are demonstrated in Fig. 9. We shall use a bar over ν, μ to indicate that the port direction of the motion is flipped like $1 \rightleftharpoons 2$ such that $\bar{\mu} = 3 - \mu$. Again denoting the energy offsets as $\epsilon_1 = \epsilon_A, \epsilon_2 = \epsilon_C, \epsilon'_1 = \epsilon_B, \epsilon'_2 = \epsilon_D$ we have the link contribution

- $[-i(\epsilon_j + \epsilon'_k)]^{-1} \delta_{kk'} \delta_{jj'} \delta_{\mu\mu'} \delta_{\nu\nu'} \quad \text{for } \overline{B_{\nu k, \mu j}} \overline{B_{\mu' j', \nu' k'}} , \quad (71)$

- $[-i(\epsilon_j + \epsilon'_k)]^{-1} \delta_{kk'} \delta_{jj'} \delta_{\bar{\mu}\mu'} \delta_{\bar{\nu}\nu'} \quad \text{for } \overline{B_{\nu k, \mu j}} \overline{B_{\nu' k', \mu' j'}} \quad \text{and} \quad \overline{B_{\mu j, \nu k}} \overline{B_{\mu' j', \nu' k''}} . \quad (72)$

The contribution of an encounter is defined by its first left port and can be written as $i(\epsilon_{j\sigma_1} + \epsilon'_{k\sigma_1})$. And again, there will be a factor $(-1)^{n_C + n_D}$ depending on the numbers of orbits included in C and D .

2. From symbol sequence to Gaussian integral

In order to represent additive terms in (68) by contractions in a Gaussian integral we interpret the symbols $B_{\nu k, \mu j}, \tilde{B}_{\mu' j', \nu' k'}$ as (Bosonic or Fermionic) variables. Again any two ports which may be connected by a link must be represented by a pair of variables which are mutually complex conjugate, up to a sign. Since rule (72) allows for contractions between two B 's differing in both their subscripts μ and ν we thus need to have

$$B_{\nu k, \mu j} = \pm B_{\bar{\nu} k, \bar{\mu} j}^* . \quad (73)$$

To write the matrices B in a more elegant way, it is helpful to divide B into four 2×2 blocks, each block having fixed direction indices μ, ν and matrix elements indexed by k, j . The diagonally opposite blocks must be mutually complex conjugate up to a sign (which may even vary with j and k). We may write

$$B \stackrel{?}{=} \begin{pmatrix} b_p & b_a^* \\ b_a & b_p^* \end{pmatrix} , \quad (74)$$

with the question mark on the equality a reminder of the as yet unspecified signs. The index p on b_p, b_p^* signals $\mu = \nu$, i.e., the ports associated with these subblocks are traversed in parallel directions by the original and the partner orbits. In contrast, the ports associated with b_a, b_a^* have $\mu \neq \nu$ and their traversal directions by original and partner orbits are antiparallel.

In addition, rule (71) allows contractions between matrix elements of B and \tilde{B} with identical subscripts. Hence these matrix elements, too, must be chosen as mutually complex conjugate (up to a sign). Since the subscripts in B and \tilde{B} are ordered differently, this means that \tilde{B} must coincide with the adjoint of B (again up to signs), i.e.,

$$\tilde{B} \stackrel{?}{=} B^\dagger . \quad (75)$$

A Gaussian-weight integral of a product of mutually conjugate matrix elements will yield a sum over all permissible ways of drawing contraction lines. To obtain the correct contribution from each link the prefactor of $B_{\mu k, \nu j} B_{\mu k, \nu j}^*$ in the exponent of the Gaussian again has to be chosen as $i(\epsilon_j + \epsilon'_k)$. Furthermore, due to the internal structure of B the sum over all $B_{\mu k, \nu j} B_{\mu k, \nu j}^*$ will include each pair of complex conjugate elements twice, which means that we have to divide the exponent by 2. Using the composite indices J, K and defining $\epsilon_J \equiv \epsilon_j$ and $\epsilon'_K \equiv \epsilon'_k$ independent of the

Greek part of J, K we can write the exponent in the Gaussian as $\sum_{JK} \pm \frac{i}{2}(\epsilon_J + \epsilon'_K) B_{KJ} B_{KJ}^*$. Incorporating also the factor $i(\epsilon_{J_{\sigma 1}} + \epsilon'_{K_{\sigma 1}})$ for each encounter we obtain

$$\pm \int d[B, \tilde{B}] e^{\sum_{J, K} \pm \frac{i}{2}(\epsilon_J + \epsilon'_K) B_{KJ} B_{KJ}^*} \\ \times \prod_{\sigma=1}^V \left(\pm i(\epsilon_{J_{\sigma 1}} + \epsilon'_{K_{\sigma 1}}) B_{K_{\sigma 1}, J_{\sigma 1}} \tilde{B}_{J_{\sigma 1}, K_{\sigma 2}} B_{K_{\sigma 2}, J_{\sigma 2}} \dots \tilde{B}_{J_{\sigma l(\sigma)}, K_{\sigma 1}} \right); \quad (76)$$

each (convergent) integral of this type includes the correct encounter and link factors for all structures with V encounters involving $l(1), l(2) \dots l(V)$ encounter stretches.

3. Signs

In order to incorporate the sign factor $(-1)^{n_C + n_D}$ and the correct prefactor of the diagonal approximation, we need to fully specify the matrices B and \tilde{B} : The off-diagonal entries in all four blocks of B are chosen Fermionic and the diagonal entries Bosonic, and the signs in (75) are picked as

$$B = \begin{pmatrix} b_p & -b_a^* \\ b_a & -\sigma_z b_p^* \end{pmatrix}, \quad \tilde{B} = \sigma_z B^\dagger; \quad (77)$$

here the multiplication with σ_z in \tilde{B} means that each 2×2 block of B^\dagger is multiplied with σ_z ; taking the adjoint of our 4×4 supermatrices involves, apart from complex conjugation and transposition in the sense of ordinary matrices, a sign flip in the lower left (Fermi-Bose) entries of all 2×2 blocks.

The signs in the Gaussian weight and the matrix product are again chosen such that both quantities can be written in terms of supertraces. The supertrace is now defined as the sum of the two upper left (Bose-Bose) entries minus the sum of the lower right (Fermi-Fermi) entries of the two diagonal blocks; in other words, the supertrace includes a negative sign for all diagonal elements associated to a Latin index 2 (regardless of the Greek index). The integral (76) then acquires the form

$$\int d[B, \tilde{B}] \exp \left(\frac{i}{2} \text{Str} \hat{\epsilon}' B \tilde{B} + \frac{i}{2} \text{Str} \hat{\epsilon} \tilde{B} B \right) (\dots) \equiv \langle\langle \dots \rangle\rangle, \quad (78)$$

with $\hat{\epsilon}, \hat{\epsilon}'$ 4×4 diagonal matrices whose diagonal elements are $\hat{\epsilon}_J = \epsilon_j$, $\hat{\epsilon}'_K = \epsilon'_k$. After eliminating all contraction lines from any integral of this type we are left with the Gaussian integral

$$\langle\langle 1 \rangle\rangle = \int d[B, \tilde{B}] \exp \left(\frac{i}{2} \text{Str} \hat{\epsilon}' B \tilde{B} + \frac{i}{2} \text{Str} \hat{\epsilon} \tilde{B} B \right) \\ = \frac{(\epsilon_1 + \epsilon'_2)^2 (\epsilon_2 + \epsilon'_1)^2}{(\epsilon_1 + \epsilon'_1)^2 (\epsilon_2 + \epsilon'_1)^2} = \frac{(\epsilon_A + \epsilon_D)^2 (\epsilon_C + \epsilon_B)^2}{(\epsilon_A + \epsilon_B)^2 (\epsilon_C + \epsilon_D)^2}. \quad (79)$$

as in the diagonal factor for time-reversal invariant systems. Therefore, upon adding to (78) the Weyl factor and the factor $1/(2^V V!)$ from (68) we obtain the contribution to the generating function (incorporating the diagonal term) as

$$Z^{(1)}(V, \{l(\sigma), J_{\sigma i}, K_{\sigma i}\}) = \frac{e^{i(\epsilon_A + \epsilon_B - \epsilon_C - \epsilon_D)/2}}{V! 2^V} \\ \times \left\langle \left\langle \prod_{\sigma=1}^V i(\epsilon_{J_{\sigma 1}} + \epsilon'_{K_{\sigma 1}}) \text{Str} P_{K_{\sigma 1}} B P_{J_{\sigma 1}} \tilde{B} P_{K_{\sigma 2}} \dots P_{J_{\sigma l(\sigma)}} \tilde{B} \right\rangle \right\rangle, \quad (80)$$

where P_J (P_K) is a projection matrix whose diagonal element associated to J (K) equals 1 while all other elements vanish.

We still have to show that with the choices made above the sign factor $(-1)^{n_C + n_D}$ for each structure is obtained correctly. The contraction rules (59) and (60) we used to prove the corresponding result in the unitary case carry over. In addition it is now possible to draw contraction lines between two B 's or two \tilde{B} 's. These lines represent links

connecting two left ports or two right ports. The corresponding contraction rules are established in Appendix A; they read

$$\left\langle\!\left\langle \text{Str}(P_K \overbrace{BP_J Y} \text{Str}(\overbrace{P_{K'} BP_{J'}} Z) \dots \right\rangle\!\right\rangle = -\frac{1}{i(\epsilon_J + \epsilon'_{K'})} \delta_{\bar{K}K'} \delta_{\bar{J}J'} \left\langle\!\left\langle \text{Str}(P_J Y P_K \tilde{Z}) \dots \right\rangle\!\right\rangle \quad (81)$$

$$\left\langle\!\left\langle \text{Str}(P_K \overbrace{BP_J Y P_{K'}} \overbrace{BP_{J'}} Z) \dots \right\rangle\!\right\rangle = -\frac{1}{i(\epsilon_J + \epsilon'_{K'})} \delta_{\bar{K}K'} \delta_{\bar{J}J'} \left\langle\!\left\langle \text{Str}(P_K \tilde{Y} P_J Z) \dots \right\rangle\!\right\rangle; \quad (82)$$

analogously for \tilde{B} . Here Y and Z are matrix products of the type required in (80) starting and ending with \tilde{B} . The sequences \tilde{Y} , \tilde{Z} are obtained from Y , Z by (i) reverting the order of matrices; (ii) replacing all B 's by \tilde{B} 's and vice versa; (iii) replacing all projectors P_K by $P_{\bar{K}}$ and P_J by $P_{\bar{J}}$. Eqs. (81) and (82) again contain the familiar factor for each link. We just chose the more compact notation

$$J = (\mu, j), \quad \bar{J} \equiv (\bar{\mu}, j), \quad K = (\nu, k), \quad \bar{K} \equiv (\bar{\nu}, k) \quad (83)$$

The Kronecker deltas in the contraction rules (81,82) imply that the connected ports must belong to the same pseudo-orbits and have opposite direction of motion.

We can now use the contraction rules (59), (60), (81) and (82) and their counterparts for contractions of two \tilde{B} 's to step by step remove the contraction lines corresponding to all links of a structure. In doing so we do not meet any sign factors apart from the steps where the final link belonging to an orbit is removed. This final link must always connect a left to a right port, since otherwise it would not be possible to close it to an orbit just by drawing an encounter stretch. Hence in this step we only have to use the rules (59) and (60) for removing links between left ports (denoted by B) and right ports (denoted by \tilde{B}). By the same argument as in Section VI we can show that the final step yields a factor $\text{Str}(P_J)$ for each orbit in $A \cup C$ and a factor $\text{Str}(P_K)$ for each orbit in $B \cup D$. According to our definition of the supertrace these factors are equal to -1 if the corresponding small indices are $j = 2$ (corresponding to the pseudo-orbit C) or $k = 2$ (corresponding to the pseudo-orbit D). We thus obtain a negative sign factor for every orbit included in C or D , and these factors combine to the desired term $(-1)^{n_C + n_D}$.

4. Towards a sigma model

Summing over structures as in Section VI and again using the assumed largeness of the imaginary part η of all energy offsets, see (16), we get the generating function

$$\begin{aligned} Z^{(1)} &\sim e^{i(\epsilon_A + \epsilon_B - \epsilon_C - \epsilon_D)/2} \int d[B, \tilde{B}] \exp \left[\frac{i}{2} \text{Str} \hat{\epsilon} \left(\sum_{l=1}^{\infty} (\tilde{B}B)^l \right) + \frac{i}{2} \text{Str} \hat{\epsilon}' \left(\sum_{l=1}^{\infty} (B\tilde{B})^l \right) \right] \\ &= \int d[B, \tilde{B}] \exp \left[\frac{i}{4} \text{Str} \left(\hat{\epsilon} \frac{1 + \tilde{B}B}{1 - \tilde{B}B} + \hat{\epsilon}' \frac{1 + B\tilde{B}}{1 - B\tilde{B}} \right) \right], \quad \eta \gg 1, \end{aligned} \quad (84)$$

equal in appearance (up to the factor $\frac{1}{2}$) to (63) for the unitary symmetry class. The matrix integral again sums up the contributions from all structures which form an asymptotic expansion in inverse powers of energy offsets. The integration domain in (84) must include the vicinity of the saddle point $B = \tilde{B} = 0$, and be such that the eigenvalues of $1 - B\tilde{B}$, $1 - \tilde{B}B$ be positive, otherwise it is arbitrary.

To see that this asymptotic series agrees with the Gaussian Orthogonal Ensemble (GOE) of RMT, we can argue like in the unitary case: The sigma model for the GOE leads to a generating function as (84) but with a specified integration domain and without the restriction $\eta \gg 1$. The large- ϵ asymptotics of the integral is determined by two saddle points of the exponent both of which have to be taken into account unless $\eta \gg 1$. The so called standard saddle at $B = 0$ leads to a contribution dominated by small B and thus coinciding with (84). The contribution of the Andreev-Altshuler saddle is related to $Z^{(1)}$ by $Z^{(2)}(\epsilon_A, \epsilon_B, \epsilon_C, \epsilon_D) = Z^{(1)}(\epsilon_A, \epsilon_B, -\epsilon_D, -\epsilon_C)$, again as in semiclassics. Our semiclassical results thus recover the asymptotics of the generating function in agreement with the sigma model.

The same must then be true for $C(\epsilon)$. The two-point correlation function of fully chaotic time-reversal invariant systems is thus asymptotically given by the sum of two divergent series in powers of ϵ^{-1} in Eq. (3).

Now we proceed to the full $C(\epsilon)$. As is well known, an analytic function can, under rather general conditions, be uniquely restored from its asymptotic series by Borel summation [6]. That method involves term-by-term Fourier transform of the asymptotic expansion leading to a converging series, followed by the inverse Fourier transform of the resulting analytic function. The first stage of the Borel summation applied to the two components of the correlation function gives the spectral form factor $K(\tau)$ both for $\tau < 1$ and $\tau > 1$. The inverse Fourier transform recovers the closed RMT expression for the correlation function in Eq. (2), for all energies ϵ , thus completing the semiclassical explanation of universality.

VIII. CONCLUSIONS

We have shown that chaotic systems without any symmetries (beyond time reversal invariance) display universal spectral statistics. To obtain both the non-oscillatory and oscillatory contributions to the spectral correlator $R(\epsilon)$ we had to (i) represent the correlator through derivatives of a generating function Z , (ii) semiclassically expand Z as a sum over quadruplets of pseudo-orbits (sets of classical periodic orbits), and (iii) uncover a Weyl symmetry of Z . The non-oscillatory and the oscillatory part of the correlation function are due to two summands of Z related by the Weyl symmetry. The oscillatory one could only be recovered in a semiclassical approximation that explicitly preserves unitarity; in this approximation the oscillations are connected to a phase factor which accounts for the average level density.

Constructively interfering contributions to the sum over pseudo-orbit quadruplets originate from orbits that are either identical or differ in close encounters in phase space. They define asymptotic $\frac{1}{\epsilon}$ -series. For dynamics without time reversal invariance those series terminate after the first term, both for the non-oscillatory and oscillatory part. For time-reversal invariant systems, the series are infinite ones but can be summed to give $R(\epsilon)$ in closed form.

To implement the summation we write the semiclassical series for the generating function Z as a matrix integral. The sum over all possible topologies then turns into an integral well known from the zero dimensional sigma model of random-matrix theory.

For both the unitary and the orthogonal symmetry class, the correlator $R(\epsilon)$ comes out as the closed-form expression familiar from random-matrix theory. In contrast to the primary $\frac{1}{\epsilon}$ -series which makes sense only for $|\epsilon| \gg 1$, the final result holds for any real value of the energy offset and yields, for $\epsilon \ll 1$, the familiar power law $R(\epsilon) \propto |\epsilon|^\beta$, with the exponent β revealing the degree of level repulsion characterizing each symmetry class. Our results thus imply the semiclassical explanation of universal level repulsion. Likewise, all other characteristics of universal spectral fluctuations based on the two-point correlator $R(\epsilon)$, like the so called spectral rigidity [36, 37], are now semiclassically understood.

The techniques here elaborated (in particular the analogy between orbit-based semiclassics and the sigma model) will be, and in fact are already helpful in tackling further challenges in the field. Examples are (i) localization effects in long thin wires [38], (ii) non-universal effects related to the Ehrenfest time (which has no counterpart in RMT) [39], (iii) a semiclassical treatment of arithmetic billiards [40] (where intrinsically quantum Hecke symmetries have strong action degeneracies as classical counterparts), (iv) spectral correlations of higher order [41] and the density of nearest-neighbor spacings, (v) transport through ballistic chaotic conductors [32, 42] including the “full counting statistics” [43], (vi) further symmetry classes, (vii) transitions between symmetry classes [44], and (viii) last but not least, the relative status of the ballistic sigma model and periodic-orbit theory [23].

Acknowledgments

We gratefully acknowledge helpful discussions with Jon Keating, Taro Nagao, Keiji Saito, Martin Sieber, Ben Simons, Hans-Jürgen Sommers, and Martin Zirnbauer, as well as funding from the Sonderforschungsbereich TR 12 of the Deutsche Forschungsgemeinschaft.

APPENDIX A: CONTRACTION RULES

In this Appendix we want to derive the contraction rules (59), (60), (81), and (82) from Eqs. (49) and (50). We start with **systems without time-reversal invariance**. First of all, contractions will be nonzero if the contracted matrix elements are mutually complex conjugate. Using the definition of the exponent according to Eq. (54) we obtain for contractions between B_{kj} and $\tilde{B}_{j'k'}$ (the latter agreeing up to the sign with $B_{k'j'}^*$),

$$\left\langle\left\langle B_{kj} \tilde{B}_{j'k'} g(B, \tilde{B}) \right\rangle\right\rangle = -\frac{s_j}{i(\epsilon_j + \epsilon_k')} \delta_{jj'} \delta_{kk'} \left\langle\left\langle g(B, \tilde{B}) \right\rangle\right\rangle. \quad (\text{A1})$$

Here s_j is a sign factor equal to 1 if $j = 1$ and to -1 if $j = 2$. To understand its origin one has to take into account the contraction rule (49) for the Bosonic variables ($j = k$) yielding $-\frac{1}{ie}$ and (50) for the Fermionic variables ($j \neq k$) yielding $\frac{1}{ie}$. An additional factor -1 emerges when $j = k = 2$ due to the negative sign in $\tilde{B}_{22} = -B_{22}^*$ and when $j = 1, k = 2$ because of minus at the term proportional to $B_{21}\tilde{B}_{12}$ in the exponent of (54).

Using Eq. (A1) we obtain the rule (59) for contractions between B and \tilde{B} in different supertraces (inter-encounter contractions) which comes about as

$$\begin{aligned}
\left\langle\left\langle \text{Str}(P_k \overbrace{BP_j Y} \text{Str}(P_{k'} \overbrace{XP_j B} \dots) \right\rangle\right\rangle &= \left\langle\left\langle \text{Str}(P_j Y \overbrace{P_k B} \text{Str}(P_{j'} \overbrace{B P_{k'} X} \dots) \right\rangle\right\rangle \\
&= \left\langle\left\langle s_j Y_{jk} \overbrace{B_{kj} s_{j'} \overbrace{B_{j'k'} X_{k'j'}}} \dots \right\rangle\right\rangle \\
&= -\frac{s_j}{i(\epsilon_j + \epsilon'_k)} \delta_{jj'} \delta_{kk'} \langle\langle s_j Y_{jk} s_j X_{kj} \dots \rangle\rangle \\
&= -\frac{1}{i(\epsilon_j + \epsilon'_k)} \delta_{jj'} \delta_{kk'} \langle\langle s_j Y_{jk} X_{kj} \dots \rangle\rangle \\
&= -\frac{1}{i(\epsilon_j + \epsilon'_k)} \delta_{jj'} \delta_{kk'} \langle\langle \text{Str}(P_j Y P_k X) \dots \rangle\rangle. \tag{A2}
\end{aligned}$$

Here we used the cyclic invariance of the supertrace, then wrote out the supertrace in components, with the help of $\text{Str}(P_k Z) = s_k Z_{kk}$ for any supermatrix Z , and finally used Eq. (A1).

Similarly Eq. (60) for intra-encounter contractions between B and \tilde{B} follows from

$$\begin{aligned}
\left\langle\left\langle \text{Str}(P_k \overbrace{BP_j U P_{j'} \overbrace{BP_k V} \dots} \right\rangle\right\rangle &= \left\langle\left\langle s_k B_{kj} \overbrace{U_{jj'} \overbrace{B_{j'k'} V_{k'k}}} \dots \right\rangle\right\rangle \\
&= \left\langle\left\langle s_k B_{kj} \overbrace{\tilde{B}_{j'k'} U_{jj'} V_{k'k}} \dots \right\rangle\right\rangle \\
&= -\frac{s_j}{i(\epsilon_j + \epsilon'_k)} \delta_{jj'} \delta_{kk'} \langle\langle s_k U_{jj} V_{kk} \dots \rangle\rangle \\
&= -\frac{1}{i(\epsilon_j + \epsilon'_k)} \delta_{jj'} \delta_{kk'} \langle\langle \text{Str}(P_j U) \text{Str}(P_k V) \dots \rangle\rangle; \tag{A3}
\end{aligned}$$

here in the second line we could interchange $U_{jj'}$ and $\tilde{B}_{j'k'}$, since a nonzero result arises only for $j = j'$ in which case U_{jj} is Bosonic and thus commutes with all other variables.

We now turn to **time-reversal invariant systems**. For these systems the derivation of rules (59) and (60) carries over directly, the only difference being that J, K are now double indices consisting of the Greek and Latin parts, $J = (\mu, j)$, $K = (\nu, k)$; the sign factor s_J has the values 1 for $j = 1$ and -1 for $j = 2$ regardless of the Greek index, and s_K is defined analogously. In addition the fact that the matrix elements of B and \tilde{B} are pairwise mutually conjugate allows to draw contraction lines between two B 's or two \tilde{B} 's. We thus obtain further contraction rules (81) and (82) for contractions between two B 's, and analogous rules for contractions between two \tilde{B} 's. To derive these contraction rules we use the important symmetry of B as defined in (77),

$$\Sigma B^t \Sigma^t = -\tilde{B} \quad \Sigma \tilde{B}^t \Sigma^t = -B \tag{A4}$$

where B^t is the transpose of B , and we have introduced $\Sigma = \begin{pmatrix} 0 & \sigma_z \\ 1 & 0 \end{pmatrix}$ and $\Sigma^t = \Sigma^{-1} = \begin{pmatrix} 0 & 1 \\ \sigma_z & 0 \end{pmatrix}$. (Readers who want to check (A4) are reminded that transposition of a supermatrix involves interchanging the indices of the matrix elements and afterwards flipping the sign of the lower left elements in each 2×2 block.) Similarly conjugating the projection matrices P_K, P_J we get

$$\Sigma P_K \Sigma^t = P_{\bar{K}}, \quad \Sigma P_J \Sigma^t = P_{\bar{J}}; \tag{A5}$$

Now let us consider the matrix products Y and Z on the l.h.s. of (81) and (82), involving alternating sequences of B 's and \tilde{B} 's and interspersed projection matrices. Under transposition and conjugation with Σ these products have (i) the order of matrices inverted (due to the transposition), (ii) B replaced by \tilde{B} and vice versa (due to (A4)), and (iii) the subscripts J and K of all projection matrices replaced by \bar{J} and \bar{K} (due to (A5)). Since the numbers of B 's and \tilde{B} 's are equal the sign factors from (A4) all mutually compensate.

Thus equipped we can prove rule (81) for inter-encounter contractions between two B 's. By using the invariance of

the supertrace both under transposition of its argument and under conjugation with Σ we write

$$\begin{aligned} \left\langle\!\left\langle \text{Str}(P_K \overline{BP_J Y}) \text{Str}(\overline{P_{K'} B} P_{J'} Z) \dots \right\rangle\!\right\rangle &= \left\langle\!\left\langle \text{Str}(P_K \overline{BP_J Y}) \text{Str}(\Sigma Z^t P_{J'} \overline{B^t P_{K'} \Sigma^t}) \dots \right\rangle\!\right\rangle \\ &= \left\langle\!\left\langle \text{Str}(P_K \overline{BP_J Y}) \text{Str}(\tilde{Z} \overline{P_{J'} B} P_{\bar{K}'}) \dots \right\rangle\!\right\rangle \\ &= -\frac{1}{i(\epsilon_J + \epsilon'_K)} \delta_{\bar{J}, J'} \delta_{\bar{K}, K'} \left\langle\!\left\langle \text{Str}(P_J Y P_K \tilde{Z}) \dots \right\rangle\!\right\rangle. \end{aligned} \quad (\text{A6})$$

Here in the third line \tilde{Z} denotes the product obtained from Z by steps (i)-(iii) above. We have thus proven Eq. (81).

We now move on to rule (82) for intra-encounter contractions between two B 's, as in

$$\left\langle\!\left\langle \text{Str}(P_K \overline{BP_J Y P_{K'} B} P_{J'} Z) \dots \right\rangle\!\right\rangle. \quad (\text{A7})$$

First of all note that the contracted matrix elements are B_{KJ} and $B_{K'J'}$, the latter coinciding up to a sign with $B_{\bar{K}, \bar{J}}^*$ (cf. Eq. (77)). Hence the only relevant case is that $J' = \bar{J}$, $K' = \bar{K}$. For this case we now write

$$\left\langle\!\left\langle \text{Str}(P_K \overline{BP_J Y P_{\bar{K}} B} P_{\bar{J}} Z) \dots \right\rangle\!\right\rangle = \left\langle\!\left\langle s_K \overline{B_{KJ} (Y P_{\bar{K}} B)}_{J\bar{J}} Z_{\bar{J}K} \dots \right\rangle\!\right\rangle \quad (\text{A8})$$

and evaluate the bracket in (A8) as

$$\begin{aligned} (Y P_{\bar{K}} B)_{J\bar{J}} &= [\Sigma^t (\Sigma B^t \Sigma^t) (\Sigma P_{\bar{K}} \Sigma^t) (\Sigma Y^t \Sigma^t) \Sigma]_{J\bar{J}}^t \\ &= (\Sigma^t \tilde{B} P_{\bar{K}} \tilde{Y} \Sigma)_{J\bar{J}} \\ &= \Sigma_{\bar{J}, J}^t \tilde{B}_{J, K} \tilde{Y}_{K, \bar{J}} \Sigma_{\bar{J}, J} \\ &= s_J \tilde{B}_{J, K} \tilde{Y}_{K, \bar{J}}. \end{aligned} \quad (\text{A9})$$

Here we first replaced the matrix sequence by itself double transposed; note that double transposition of a supermatrix (not an identity operation!) leaves its Bosonic elements like the one with the subscripts $J\bar{J}$ unchanged. Then we invoked Eqs. (A4,A5) and noted that for the matrix elements at hand transposition just amounts to interchanging the two subscripts. In the third line we used that only those elements of Σ, Σ^t are nonzero for which one subscript is equal to another one barred, i.e., the port directions of motion are opposite. Finally, we used $\Sigma_{\bar{J}, J}^t \Sigma_{\bar{J}, J} = s_J$. Inserting (A9) into (A8) we now obtain

$$\begin{aligned} \left\langle\!\left\langle \text{Str}(P_K \overline{BP_J Y P_{\bar{K}} B} P_{\bar{J}} Z) \dots \right\rangle\!\right\rangle &= \delta_{\bar{J}, J'} \delta_{\bar{K}, K'} \left\langle\!\left\langle s_K \overline{B_{KJ} s_J \tilde{B}_{J, K} \tilde{Y}_{K, \bar{J}} Z_{\bar{J}, K}} \dots \right\rangle\!\right\rangle \\ &= -\frac{1}{i(\epsilon_J + \epsilon'_K)} \delta_{\bar{J}, J'} \delta_{\bar{K}, K'} \left\langle\!\left\langle s_K \tilde{Y}_{K, \bar{J}} Z_{\bar{J}, K} \dots \right\rangle\!\right\rangle \\ &= -\frac{1}{i(\epsilon_J + \epsilon'_K)} \delta_{\bar{J}, J'} \delta_{\bar{K}, K'} \left\langle\!\left\langle \text{Str}(P_K \tilde{Y} P_{\bar{J}} Z) \dots \right\rangle\!\right\rangle \end{aligned} \quad (\text{A10})$$

which is (82). Here we used (A1), defined \tilde{Y} in analogy to \tilde{Z} , and in the end rearranged the result as a supertrace.

One easily shows that rules analogous to (81) and (82) also hold for contractions between two \tilde{B} 's.

APPENDIX B: SUPERSYMMETRIC SIGMA MODEL

We propose to briefly review the basics of the sigma model; details can be found in [3, 4, 14, 31, 33, 34]. The sigma model is a way to implement averages over random matrices. The appropriate random-matrix ensemble for systems without time-reversal invariance is the Gaussian Unitary Ensemble (GUE), consisting of complex Hermitian $N \times N$ matrices H with a Gaussian weight proportional to $e^{-\frac{N}{2} \text{tr } H^2}$; the matrix dimension N is sent to infinity. The pertinent generating function Z looks like (5), but with the energy average $\langle \dots \rangle$ replaced by the GUE average. To perform that average, three variants of the sigma model are available, all based on representing determinants by suitable Gaussian integrals. The supersymmetric variant deals with superintegrals over both ordinary and Grassmannian variables while two replica variants exclusively use either Grassmann or ordinary Gaussian integrals.

1. From GUE to supermatrix integral over saddle-point manifold

In the supersymmetric approach each of the determinants in the generating function (5) is represented by a Gaussian integral, Fermionic for the numerator and Bosonic for the denominator,

$$\begin{aligned} Z(\epsilon_A, \epsilon_B, \epsilon_C, \epsilon_D) &= (-1)^N \left(\prod_{m=1}^N \int \frac{d^2 z_{Am}}{\pi} \frac{d^2 z_{Bm}}{\pi} d\eta_{Cm}^* d\eta_{Cm} d\eta_{Dm}^* d\eta_{Dm} \right) \\ &\times \left\langle e^{-i \sum_{m,n} \{ -z_{Am}^*(E + \epsilon_A - H)_{mn} z_{An} + z_{Bm}^*(E - \epsilon_B - H)_{mn} z_{Bn} \}} \right. \\ &\quad \left. \times e^{-i \sum_{m,n} \{ \eta_{Cm}^*(E + \epsilon_C - H)_{mn} \eta_{Cn} + \eta_{Dm}^*(E - \epsilon_D - H)_{mn} \eta_{Dn} \}} \right\rangle; \end{aligned} \quad (\text{B1})$$

at the expense of the overall factor $(-1)^N$ the signs of the ‘‘Bosonic’’ quadratic forms $z_A^*(\dots)z_A$ and $z_B^*(\dots)z_B$ have been chosen so as to make the Bosonic Gaussian integrals convergent, in consistency with $\text{Im } \epsilon_{A,B,C,D} > 0$. $\eta_C, \eta_C^*, \eta_D, \eta_D^*$ are anticommuting Grassmann variables. As usual, the asterisk on commuting variables means complex conjugation. On anticommuting variables the asterisk may, for the purposes of the present paper, be given the same interpretation (complemented by $(\eta^*)^* = -\eta$, see [3]).

At this point hooking onto the presentation in Chapt. 10 of Ref. [3]⁷, we introduce the supervector

$$\Phi = \begin{pmatrix} \Phi_1 \\ \Phi_2 \\ \Phi_3 \\ \Phi_4 \end{pmatrix} = \begin{pmatrix} z_B \\ \eta_D \\ z_A \\ \eta_C \end{pmatrix} \quad (\text{B2})$$

where each of the four Φ_α ’s has N components $\Phi_{\alpha m}$; the ordering of the Φ_α in (B2) is known as the advanced-retarded notation. To accommodate the relative sign of the Bosonic quadratic forms mentioned and the different energy arguments above we employ the (4×4) matrices,

$$\begin{aligned} L &= \text{diag}(1, 1, -1, 1) = \begin{pmatrix} 1 & 0 \\ 0 & -\sigma_z \end{pmatrix}, \\ \hat{E} &= E - \frac{1}{2\pi\rho} \text{diag}(\epsilon_B, \epsilon_D, -\epsilon_A, -\epsilon_C) \equiv E - \frac{1}{2\pi\rho} \begin{pmatrix} \hat{\epsilon}' & 0 \\ 0 & -\hat{\epsilon} \end{pmatrix} \end{aligned} \quad (\text{B3})$$

whose block form will be convenient in the future. Our generating function may then be succinctly written as

$$Z = (-1)^N \int d\Phi^* d\Phi \left\langle \exp \left\{ i\Phi^\dagger L [H - \hat{E}] \Phi \right\} \right\rangle, \quad (\text{B4})$$

with $d\Phi^* d\Phi$ denoting the integration measure written explicitly in (B1). The Gaussian average over Hamiltonian matrices H with a normalized Gaussian weight $\propto e^{-\frac{N}{2}\text{tr}H^2}$ leads to $Z = (-1)^N \exp \left(-\frac{1}{2N} \sum_{\alpha,\beta}^{1\dots 4} L_\alpha L_\beta \sum_{m,n}^{1\dots N} \Phi_{\alpha m}^* \Phi_{\alpha n} \Phi_{\beta n}^* \Phi_{\beta m} \right)$. The Hermitian 4×4 supermatrix

$$\tilde{Q}_{\alpha\beta} = \frac{1}{N} \sum_{m=1}^N \Phi_{\alpha m} \Phi_{\beta m}^* \quad (\text{B5})$$

allows to further compact Z as

$$Z = (-1)^N \int d\Phi^* d\Phi \exp \left\{ N \text{Str} \left[-i\hat{E}\tilde{Q}L - \frac{1}{2}(\tilde{Q}L)^2 \right] \right\} \quad (\text{B6})$$

where the symbol Str denotes the supertrace. The integrand would be invariant under the transformation $\tilde{Q} \rightarrow T\tilde{Q}T^\dagger$ with pseudounitary matrices T defined by

$$T^\dagger LT = L, \quad (\text{B7})$$

⁷ We follow the notation of [3] apart from reordering the elements of Φ and setting the parameter λ in the random-matrix integrals equal to 1.

if the (real part of) the energy matrix were proportional to the unit matrix. Due to

$$\bar{\rho}(E=0) = N/\pi \quad (\text{B8})$$

the energy offsets $\epsilon_{A,B,C,D}/2\pi\bar{\rho}$ are of only order $1/N$, which means that in any case this symmetry persists to leading order in $1/N$.

The exponent in (B6) is quartic in the integration variables Φ, Φ^* . Upon employing a suitable Hubbard-Stratonovich transformation [3] we render the Φ -integral Gaussian and end up with an integral over 4×4 supermatrices Q ,

$$Z = - \int dQ \exp \left\{ -\frac{N}{2} \text{Str} \left[Q^2 + 2 \ln(\hat{E} - Q) \right] \right\}, \quad (\text{B9})$$

with a flat integration measure dQ . For a details of the transformation, in particular the overall minus sign and the integration range in (B9), we refer the reader to Ref. [3]. Important for what follows is that the integration range is restricted to matrices Q diagonalizable by pseudounitary transformations T .

The large- N limit allows for a saddle-point approximation. The saddle manifold is determined by

$$Q = \frac{1}{\hat{E} - Q}. \quad (\text{B10})$$

We may confine the discussion of this equation to $E = 0$, which means that we evaluate Z in the middle of Wigner's semicircle. As seen the energy offsets $\hat{E} - E$ are of order $1/N$ and thus irrelevant for saddles, we may even drop \hat{E} entirely. The saddle-point equation then simplifies to $Q^2 = -1$.

Next, we expand the exponent in Eq. (B9) about the saddle manifold $Q^2 = -1$, to first order in the energy offsets and still sticking to the center of the Wigner's semicircle, $E = 0$, such that $\hat{E} = -\frac{1}{2\pi\bar{\rho}} \begin{pmatrix} \epsilon' & 0 \\ 0 & -\epsilon \end{pmatrix}$. Since the supertrace of the unit matrix vanishes we get $\text{Str}[Q^2 + 2 \ln(\hat{E} - Q)] = \text{Str}[\ln(\hat{E}^2 - 2\hat{E}Q + Q^2)] \approx \text{Str}[\ln(-1) + \ln(1 + 2\hat{E}Q)] \approx \text{Str}2\hat{E}Q$, where we again neglected corrections of higher order in $1/N$. Going to the limit $N \rightarrow \infty$ we get the integral

$$Z = - \int_{Q^2=-1} dQ e^{\frac{1}{2} \text{Str} Q \begin{pmatrix} \epsilon' & 0 \\ 0 & -\epsilon \end{pmatrix}}. \quad (\text{B11})$$

The eigenvalues of the matrices Q must be equal to $\pm i$. This means that the saddle-point manifold can be divided into 5 conjugacy classes differing by the numbers n_+, n_- of the eigenvalues $+i$ and $-i$. In fact, only one of them is of relevance [3], with $n_+ = n_- = 2$. Since the matrices Q must be diagonalizable by pseudo-unitary transformations the relevant Q 's can thus be represented as

$$Q = i T \Lambda^{(1)} T^{-1} \quad (\text{B12})$$

with the diagonal matrix

$$\Lambda^{(1)} = \text{diag}(1, 1, -1, -1). \quad (\text{B13})$$

2. Rational parametrization

There is an ambiguity in the choice of the matrices $T = \begin{pmatrix} T_{aa} & T_{ar} \\ T_{ra} & T_{rr} \end{pmatrix}$. Due to the degeneracy of the eigenvalues of $\Lambda^{(1)}$, Q remains unchanged if we replace T by the product TK with K any invertible block diagonal matrix. Choosing $K = \begin{pmatrix} T_{aa}^{-1} & 0 \\ 0 & T_{rr}^{-1} \end{pmatrix}$ we have

$$TK = 1 + \begin{pmatrix} 0 & B \\ \tilde{B} & 0 \end{pmatrix}, \quad (TK)^{-1} = \begin{pmatrix} (1 - B\tilde{B})^{-1} & -B(1 - \tilde{B}B)^{-1} \\ -\tilde{B}(1 - B\tilde{B})^{-1} & (1 - \tilde{B}B)^{-1} \end{pmatrix} \quad (\text{B14})$$

with $B = T_{ar}T_{rr}^{-1}$ and $\tilde{B} = T_{ra}T_{aa}^{-1}$. The pseudounitarity of T entails the matrix $(TK)^\dagger L(TK) = K^\dagger L K = \begin{pmatrix} (T_{aa}^{-1})^\dagger (T_{aa}^{-1}) & 0 \\ 0 & -(T_{rr}^{-1})^\dagger \sigma_z (T_{rr}^{-1}) \end{pmatrix}$ to be block diagonal. On the other hand

$$(TK)^\dagger L(TK) = \begin{pmatrix} 1 - \tilde{B}^\dagger \sigma_z \tilde{B} & B - \tilde{B}^\dagger \sigma_z \\ B^\dagger - \sigma_z \tilde{B} & B^\dagger B - \sigma_z \end{pmatrix}, \quad (\text{B15})$$

such that block diagonality yields $\tilde{B} = \sigma_z B^\dagger$. With the latter relation between B and \tilde{B} established we can write these 2×2 supermatrices as

$$B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}, \quad B^\dagger = \begin{pmatrix} B_{11}^* & B_{21}^* \\ -B_{12}^* & B_{22}^* \end{pmatrix}, \quad \tilde{B} = \sigma_z B^\dagger = \begin{pmatrix} B_{11}^* & B_{21}^* \\ B_{12}^* & -B_{22}^* \end{pmatrix}; \quad (\text{B16})$$

note that the diagonal elements B_{11}, B_{22} are Bosonic (commuting) variables whereas the off-diagonals B_{12}, B_{21} are Fermionic, in accordance with the advanced-retarded representation we are using. From this point on we rename TK as T .

The saddle-point manifold results as

$$T\Lambda^{(1)}T^{-1} = \begin{pmatrix} \frac{1+B\tilde{B}}{1-BB} & -2B(1-\tilde{B}B)^{-1} \\ 2\tilde{B}(1-B\tilde{B})^{-1} & -\frac{1+\tilde{B}B}{1-BB} \end{pmatrix}. \quad (\text{B17})$$

The final integral over the manifold becomes an integral over the 2×2 matrices B, \tilde{B} ,

$$Z \sim - \int d[B, \tilde{B}] \exp \left[\frac{i}{2} \text{Str} T\Lambda^{(1)}T^{-1} \begin{pmatrix} \hat{\epsilon}' & 0 \\ 0 & -\hat{\epsilon} \end{pmatrix} \right] \quad (\text{B18})$$

$$= - \int d[B, \tilde{B}] \exp \left[\frac{i}{2} \text{Str} \left(\hat{\epsilon} \frac{1+\tilde{B}B}{1-B\tilde{B}} + \hat{\epsilon}' \frac{1+B\tilde{B}}{1-B\tilde{B}} \right) \right] \\ = - e^{\frac{i}{2}(\epsilon_A + \epsilon_B - \epsilon_C - \epsilon_D)/2} \int d[B, \tilde{B}] \exp \left[i \text{Str} \left(\hat{\epsilon} \frac{\tilde{B}B}{1-B\tilde{B}} + \hat{\epsilon}' \frac{B\tilde{B}}{1-B\tilde{B}} \right) \right]. \quad (\text{B19})$$

with the flat measure $d[B, B] = \frac{d^2 B_{11}}{\pi} \frac{d^2 B_{22}}{\pi} dB_{12} * dB_{12} dB_{21}^* dB_{21}$. The integration ranges for the Bosonic integrals over B_{11} and B_{22} are respectively $|B_{11}|^2 < 1$ and $|B_{22}|^2 < \infty$. Note that the foregoing integral representation of $Z^{(1)}$ equals the one established semiclassically, see (63).

We now proceed to the asymptotics of the remaining integral (B18) in the limit of large energy offsets, $|\epsilon_{A,B,C,D}| \gg 1$. To that end we employ one more saddle-point approximation⁸. The only finite stationary point of the exponent in (B18) occurs at $B = \tilde{B} = 0$. Let us assume, momentarily, that the energy offsets are provided with a large positive imaginary part, $\eta \gg 1$. Then this will be the only critical point of the integrand responsible for the high energy asymptotics of the integral (B18). To get an asymptotic series of Z in powers of ϵ^{-1} we expand the matrix $(1-B\tilde{B})^{-1}$ and $(1-\tilde{B}B)^{-1}$ in the exponent in powers of B, \tilde{B} about the unit matrix and further expand the resulting exponential about its Gaussian part. Scaling the integration variables as $B = B'/\sqrt{|\epsilon|}$ and similarly for \tilde{B} with a typical dimensionless energy offset ϵ close to $\epsilon_{A,B,C,D}$ we see that the higher-order terms in the expansion indeed become small perturbations in the high-energy limit. The Gaussian superintegral giving the leading-order contribution reads

$$Z^{(1)} \sim - e^{\frac{i}{2}(\epsilon_A + \epsilon_B - \epsilon_C - \epsilon_D)/2} \\ \times \int \frac{d^2 B_{11}}{\pi} \frac{d^2 B_{22}}{\pi} e^{i\{B_{11}^* B_{11}(\epsilon_A + \epsilon_B) + B_{22} B_{22}^*(\epsilon_C + \epsilon_D)\}} \\ \times \int dB_{12}^* B_{12} dB_{21}^* dB_{21} e^{i\{B_{12}^* B_{12}(\epsilon_B + \epsilon_C) + B_{21}^* B_{21}(\epsilon_A + \epsilon_D)\}}. \quad (\text{B20})$$

The integration domain for the Bosonic variables can be extended to the full complex plane; in view of $\eta \gg 1$ the error thus introduced is exponentially small. Upon doing the remaining elementary integrals we find

$$Z(\epsilon_A, \epsilon_B, \epsilon_C, \epsilon_D) \sim Z^{(1)}(\epsilon_A, \epsilon_B, \epsilon_C, \epsilon_D) \sim e^{i(\epsilon_A + \epsilon_B - \epsilon_C - \epsilon_D)/2} \frac{(\epsilon_A + \epsilon_D)(\epsilon_B + \epsilon_C)}{(\epsilon_A + \epsilon_B)(\epsilon_C + \epsilon_D)}, \quad \eta \gg 1. \quad (\text{B21})$$

⁸ The previous saddle-point approximation had the matrix dimension N as its large parameter and we could take $\epsilon_{A,B,C,D}/N \rightarrow 0$ as $N \rightarrow \infty$. No conflict arises with subsequently considering the large- ϵ asymptotics.

3. Evaluation of the matrix integral

In fact the result (B21) fully exhausts $Z^{(1)}$: In agreement with the Duistermaat-Heckmann theorem [45], all further terms of the perturbation series vanish for the unitary symmetry class; in the framework of the semiclassical theory this was already demonstrated in Section V.

To prove the absence of perturbative corrections in the present RMT context we transform integration variables such that the integral in (B18) becomes Gaussian and similar to (B20). To that end we introduce an analogue of the singular-value decomposition of the matrix B ,

$$B = U \underline{B} V^{-1}, \quad \underline{B} = \text{diag}(x_B, x_F) \quad (\text{B22})$$

with unitary U and pseudo-unitary V , and parametrize the latter matrices as

$$U = \begin{pmatrix} 1 + \eta\eta^*/2 & \eta \\ \eta^* & 1 - \eta\eta^*/2 \end{pmatrix} \begin{pmatrix} e^{-i\phi_B} & 0 \\ 0 & e^{-i\phi_F} \end{pmatrix}, \quad V = \begin{pmatrix} 1 - \tau\tau^*/2 & \tau \\ -\tau^* & 1 + \tau\tau^*/2 \end{pmatrix} = V^\dagger = \sigma_z V^{-1} \sigma_z. \quad (\text{B23})$$

The matrix \underline{B} is diagonal with positive (numerical parts of the) diagonal entries, with the Bose-Bose entry obeying $0 \leq x_B < 1$, due to $|B_{11}| < 1$. Obviously, $\tilde{B} = V \sigma_z \underline{B} U^{-1} \equiv V \tilde{B} U^{-1} = \text{diag}(x_B, -x_F)$, with the Fermi-Fermi entry negative. It further follows that the matrix products $B\tilde{B}$ and $\tilde{B}B$ are respectively diagonalized by U and V ,

$$U^{-1} B \tilde{B} U = V^{-1} \tilde{B} B V = \underline{B} \tilde{B} = \begin{pmatrix} x_B^2 & 0 \\ 0 & -x_F^2 \end{pmatrix} \equiv \text{diag}(l_B, l_F), \quad (\text{B24})$$

with the eigenvalues $0 \leq l_B < 1$, $l_F \leq 0$. The matrices U, V also diagonalize the rational functions of $B\tilde{B}$ and $\tilde{B}B$ appearing in the sigma-model integral (63),

$$U^{-1} \frac{B\tilde{B}}{1 - B\tilde{B}} U = V^{-1} \frac{\tilde{B}B}{1 - \tilde{B}B} V = \text{diag} \left(\frac{l_B}{1 - l_B}, \frac{l_F}{1 - l_F} \right) \equiv \text{diag}(m_B, m_F). \quad (\text{B25})$$

The representation (B22) can be inverted to express the variables on the right-hand side in terms of the original elements of B, \tilde{B} . The uniqueness of the decomposition of our B is thus ascertained. Moreover, the eight new variables $l_B, l_F, \phi_B, \phi_F, \eta^*, \eta, \tau^*, \tau$ can be employed as integration variables instead of the original elements of B, \tilde{B} . A somewhat tedious calculation of the Berezinian yields the measure

$$d[B, \tilde{B}] = \frac{dl_B dl_F d\phi_B d\phi_F d\eta^* d\eta d\tau^* d\tau}{4\pi^2 (l_B - l_F)^2}. \quad (\text{B26})$$

Our next step is yet another change of integration variables which will be elements of matrices analogous to B, \tilde{B} but with $l_{B,F}$ replaced by $m_{B,F}$,

$$C = U \text{diag}(\sqrt{m_B}, \sqrt{-m_F}) V^{-1}, \quad \tilde{C} = \sigma_z C^\dagger. \quad (\text{B27})$$

We obviously have

$$C \tilde{C} = \frac{B\tilde{B}}{1 - B\tilde{B}}, \quad \tilde{C} C = \frac{\tilde{B}B}{1 - \tilde{B}B}, \quad (\text{B28})$$

such that the integrand in the sigma-model integral becomes Gaussian in terms of the matrices C and \tilde{C} . Moreover, in complete analogy with (B26) the integration measure can be written as

$$d[C, \tilde{C}] = \frac{dm_B dm_F d\phi_B d\phi_F d\eta^* d\eta d\tau^* d\tau}{4\pi^2 (m_B - m_F)^2}. \quad (\text{B29})$$

Comparing the measures (B26) and (B29) we get

$$d[C, \tilde{C}] = d[B, \tilde{B}] \frac{(l_B - l_F)^2}{(m_B - m_F)^2} \frac{dm_B}{dl_B} \frac{dm_F}{dl_F} = d[B, \tilde{B}]. \quad (\text{B30})$$

Therefore, in terms of the integration variables C_{ik}, C_{ik}^* the sigma-model integral becomes Gaussian. Again assuming that the energy offsets are all provided with the large positive imaginary part η we can extend the integration domain for the Bosonic variables to the full complex plane, at a risk only of corrections of the order $\exp(-\eta)$. Then the resulting C -integral will become identical with the leading-order integral (B20). We have thus proven that contribution of all higher-order terms in the expansion of the integrand about the saddle point $B = 0$ cancels, and the asymptotics of the generating function, in all orders of $\frac{1}{\epsilon}$, is exhausted (for $\eta \gg 1$) by (B21).

4. Andreev-Altshuler saddle point

If the imaginary part η of the energy arguments is small or zero $Z^{(1)}$ does not capture the full high-energy asymptotics of the generating function: an additive component $Z^{(2)}$ appears due to another, the Andreev-Altshuler or “anomalous” stationary point B^{AA} . To show it let us use the parametrization (B22) introduced in the previous subsection; note that the elements of the diagonal matrix \underline{B} satisfy $0 \leq x_B < 1$, $0 \leq x_F < \infty$. Using (B24),(B28) we can rewrite the supertrace in the exponent of the sigma-model integral like

$$\text{Str} \left(\hat{\epsilon} \frac{1 + \tilde{B}B}{1 - \tilde{B}B} + \hat{\epsilon}' \frac{1 + B\tilde{B}}{1 - B\tilde{B}} \right) = (V^{-1}\hat{\epsilon}V + U^{-1}\hat{\epsilon}'U)_{BB} \frac{1 + x_B^2}{1 - x_B^2} - (V^{-1}\hat{\epsilon}V + U^{-1}\hat{\epsilon}'U)_{FF} \frac{1 - x_F^2}{1 + x_F^2}. \quad (\text{B31})$$

At a stationary point the derivatives of this expression by x_B and x_F must be zero, from which it follows that x_B must be zero, whereas x_F can be zero or infinity. Consequently there exist two stationary points in the matrix space, (a) the standard saddle $B = 0$ giving birth to $Z^{(1)}$, and (b) the so called Andreev-Altshuler saddle $B^{\text{AA}} = \text{diag}(0, \infty_C)$ where ∞_C stands for the infinitely remote point of the complex plane. The anomalous saddle is responsible for the part $Z^{(2)}$ of the generating function exponentially small when $\eta \gg 1$ but of the same order as $Z^{(1)}$ if the energy offsets are real.

An elegant way to find the contribution of the anomalous stationary point involves retracing the steps leading to the rational parametrization and finding the point Q^{AA} of the manifold (B12) associated with B^{AA} . Let us represent B^{AA} and \tilde{B}^{AA} as limits $B^{\text{AA}} = \lim_{a \rightarrow 0} PK_a^{-1}$, $\tilde{B}^{\text{AA}} = \lim_{a \rightarrow 0} PK_{-a}^{-1}$ with the diagonal 2×2 matrices

$$P = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad K_{\pm a} = \begin{pmatrix} 1 & 0 \\ 0 & \pm a \end{pmatrix}; \quad (\text{B32})$$

for simplicity we take a real, i.e., we approach the infinite point of the complex plane along the real axis. When a is finite the respective 4×4 matrix $T = \begin{pmatrix} 1 & B \\ \tilde{B} & 1 \end{pmatrix}$ factorizes,

$$T_a = \begin{pmatrix} K_{-a} & P \\ P & K_a \end{pmatrix} \begin{pmatrix} K_{-a}^{-1} & 0 \\ 0 & K_a^{-1} \end{pmatrix}. \quad (\text{B33})$$

Since the second factor in T_a commutes with $\Lambda^{(1)}$ we find

$$\begin{aligned} Q^{\text{AA}} &= i \lim_{a \rightarrow 0} T_a \Lambda^{(1)} T_a^{-1} = iR \Lambda^{(1)} R^{-1}, \\ R = R^{-1} &= \begin{pmatrix} K_0 & P \\ P & K_0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \end{aligned} \quad (\text{B34})$$

The matrix R describes a permutation transposing the second and fourth entries, i.e., the Fermi-Fermi ones. We thus have

$$\Lambda^{(2)} \equiv R \Lambda^{(1)} R^{-1} = \text{diag}(1, -1, -1, 1). \quad (\text{B35})$$

To get the part of the high-energy asymptotics connected with the Andreev-Altshuler stationary point we reparameterize the integration variables in (B11) as $Q' = RQR$. The Jacobian of the transformation $Q \rightarrow Q'$ is unity. Denoting by \hat{E}' the offset matrix with the Fermi entries transposed, $\hat{E}' \equiv R\hat{E}R = \text{diag}(\epsilon_B, -\epsilon_C, -\epsilon_A, \epsilon_D)$ we bring the Q -integral to an equivalent form $Z = - \int_{Q'^2=-1} dQ' e^{\frac{1}{2}\text{Str} Q' \hat{E}'}$. Now we can set $Q' = T \Lambda^{(1)} T^{-1}$ and rationally parameterize T which leads to a superintegral differing from (B18) by the interchange $\epsilon_D \leftrightarrow -\epsilon_C$. Expanding the exponent around the stationary point $B = \tilde{B} = 0$ corresponding to $Q' = i\Lambda^{(1)}$, or $Q = Q^{\text{AA}} = i\Lambda^{(2)}$ we obtain an asymptotic series in ϵ^{-1} which truncates on its first term, $Z^{(2)} = Z^{(1)}|_{\epsilon_D \leftrightarrow -\epsilon_C}$. Adding the contributions of both stationary points we get the full GUE generating function $Z = Z^{(1)} + Z^{(2)}$.

APPENDIX C: BOSONIC REPLICAS

The main technical difficulty in establishing a one-to-one correspondence between semiclassics and the signal model of RMT are the spectral determinants in the numerator of the generating functions. These determinants lead

to unwelcome sign factors $(-1)^{n_C+n_D}$ in the semiclassical treatment, and in RMT they cannot be expressed through Gaussian integrals over complex numbers. In the supersymmetric treatment the way around these difficulties is to introduce anticommuting Fermionic variables. These variables incorporate the appropriate signs in semiclassics, and allow for an integral representation of non-inverted determinants in RMT. In the present appendix we want to present an alternative solution based on the so-called replica trick. (In Appendix D we will consider a combination of the replica trick with Fermionic variables.)

While the replica trick offers considerable technical convenience, there is a price to pay: First, the replica trick is hard to justify mathematically; second, at least in the present (Bosonic) variant only the part $Z^{(1)}$ is accessible. That price has let replicas fall into some disrepute. From a semiclassical point of view, the price looks more affordable. We have seen $Z^{(1)}$ to yield the full two-point correlator with the help of the “crosswise” procedure of Sect. II A or, equivalently, $Z^{(2)}$ to be accessible from $Z^{(1)}$ by the Riemann-Siegel lookalike of Sect. II C.

1. Semiclassical construction of a sigma model

To construct a replica sigma model from semiclassics, we start from formula (37). We assume that the pseudo-orbits C and D are each divided into $r-1$ subsets yielding altogether $2r$ pseudo-orbits. We denote the pseudo-orbit A by $j=1$, and the $r-1$ parts of C by $j=2, \dots, r$. Similarly B is indicated by $k=1$, whereas the $r-1$ parts of D are labelled by $k=2, \dots, r$. Each of the n_C orbits in C can be included in any of the $r-1$ subsets of C , and similarly for D . This leaves altogether $(r-1)^{n_C+n_D}$ ways for distributing the orbits of C and D among subsets. If we now formally take the limit $r \rightarrow 0$, this number of possibilities turns into the factor $(-1)^{n_C+n_D}$ needed in our semiclassical treatment. We can thus drop the sign factor if we write $Z_{\text{off}}^{(1)}$ as

$$\begin{aligned} Z_{\text{off}}^{(1)} &= \lim_{r \rightarrow 0} Z_{\text{off}}^{(1)}(r) \\ Z_{\text{off}}^{(1)}(r) &= \sum_{\substack{\text{structures} \\ \text{of } 2r-\text{tuples}}} \frac{1}{V!} \frac{\prod_{\text{enc}} i(\epsilon_j + \epsilon'_k)}{\prod_{\text{links}} (-i(\epsilon_j + \epsilon'_k))} . \end{aligned} \quad (\text{C1})$$

where we sum over structures of $2r$ -tuples rather than quadruplets of pseudo-orbits. This means that now every way of distributing the orbits in C and D among the subsets is considered to give rise to a separate structure. The indices of ϵ_j, ϵ'_k in each link factor are chosen according to the pseudo-orbits j, k that the link belongs to; the corresponding indices in the encounter factor are chosen according to the pseudo-orbits the first entrance port belongs to. In line with the above definitions of j, k we take

$$\begin{aligned} \epsilon_1 &= \epsilon_A, & \epsilon_2 &= \dots = \epsilon_r = \epsilon_C, \\ \epsilon'_1 &= \epsilon_B, & \epsilon'_2 &= \dots = \epsilon'_r = \epsilon_D . \end{aligned} \quad (\text{C2})$$

For the diagonal approximation we write

$$\begin{aligned} Z_{\text{diag}}^{(1)} &= \lim_{r \rightarrow 0} Z_{\text{off}}^{(1)}(r) \\ Z_{\text{diag}}^{(1)}(r) &= e^{i(\epsilon_A + \epsilon_B + (r-1)\epsilon_C + (r-1)\epsilon_D)/2} (-i(\epsilon_A + \epsilon_B))^{-1} (-i(\epsilon_A + \epsilon_D))^{-(r-1)} (-i(\epsilon_C + \epsilon_B))^{-(r-1)} (-i(\epsilon_C + \epsilon_D))^{-(r-1)} \\ &= e^{i \sum_j \epsilon_j / 2 + i \sum_k \epsilon'_k / 2} \prod_{j,k=1 \dots r} (-i(\epsilon_j + \epsilon'_k))^{-1} \end{aligned} \quad (\text{C3})$$

where the factors $-i$ are just inserted for later convenience; they cancel mutually in the replica limit $r \rightarrow 0$. Now we can construct a sigma model similarly as in the main part: Each entrance port belonging to an original pseudo-orbit j (either A or a part of C) and a partner pseudo-orbit k (either B or a part of D) is indicated by a symbol B_{kj} . Correspondingly each exit port is denoted by a symbol \tilde{B}_{kj} . The indices of B 's and \tilde{B} 's belonging to the same encounter have to coincide like the subscripts in the trace of a matrix product. We can again sum over all ways of connecting these ports by links with the help of a Gaussian integral over B_{kj}, \tilde{B}_{kj} . This is done in the same way as in the supersymmetric approach, with only one difference: Since the sign factor $(-1)^{n_C+n_D}$ has been eliminated by the replica trick, all sign factors in the previous treatment can be omitted including those in the commutation relation of B, \tilde{B} . The coefficients B_{kj}, \tilde{B}_{kj} can thus be taken as ordinary complex variables with $\tilde{B}_{kj} = B_{kj}^*$ (i.e. $\tilde{B} = B^\dagger$). All supermatrices are thus replaced by ordinary matrices which are now larger due to the r different values of j and k .

To go through the argument in detail, we split the sum over structures into sums over the possible numbers and sizes of encounters, assignments of ports to pseudo-orbits, and ways of connecting the ports through links:

$$\sum_{\text{structures}} = \sum_{\# \text{enc}} \sum_{\text{ports}} \sum_{\text{links}} . \quad (\text{C4})$$

The three subsums are done step by step.

Link sum: We start by summing over link connections for fixed encounters and assignment of ports to pseudo-orbits. We denote by L_{kj}^{in} , L_{kj}^{out} the numbers of entrance and exit ports forming part of the original pseudo-orbits $j = 1, 2, \dots, r$ and the partner pseudo-orbits $k = 1, 2, \dots, r$. The links now have to connect each of the L_{kj}^{out} exit ports of j, k to one of the L_{kj}^{in} corresponding entrance ports, and we can conclude $L_{kj}^{\text{in}} = L_{kj}^{\text{out}}$ for all j and k . With that condition satisfied, all connections between entrance and exit ports associated with the same j and k are permissible, i.e., there are $\prod_{jk} L_{kj}^{\text{in}}! = \prod_{jk} L_{kj}^{\text{out}}!$ possibilities. For each j, k we obtain a combinatorial factor $\delta_{L_{kj}^{\text{in}}, L_{kj}^{\text{out}}} L_{kj}^{\text{in}}!$.

We now combine the latter combinatorial factor with the $L_{kj} + 1$ contributions $\frac{1}{-i(\epsilon_j + \epsilon'_k)}$, one from each link within j and k and one from the diagonal approximation. We can represent the product as an integral over complex parameters B_{kj} , where we let $\tilde{B}_{jk} = B_{kj}^*$,

$$(-i(\epsilon_j + \epsilon'_k))^{-(L_{kj}^{\text{in}} + 1)} \delta_{L_{kj}^{\text{in}}, L_{kj}^{\text{out}}} L_{kj}^{\text{in}}! = \int d^2 B_{kj} e^{i(\epsilon_j + \epsilon'_k)|B_{kj}|^2} (B_{kj})^{L_{kj}^{\text{in}}} (\tilde{B}_{jk})^{L_{kj}^{\text{out}}} \quad (\text{C5})$$

with $d^2 B_{kj} = \frac{d\text{Re}B_{kj} d\text{Im}B_{kj}}{\pi}$, for each pair j and k . The integral converges due to $\text{Im} \epsilon_j, \text{Im} \epsilon'_k > 0$.

The only terms not yet taken into account are the encounter contributions $i(\epsilon_j + \epsilon'_k)$, and the factor $e^{i(\sum_j \epsilon_j + \sum_k \epsilon'_k)/2} \frac{1}{V!}$ from Eqs. (C1) and (C3). With these terms included the link sum yields the generating function

$$Z^{(1)}(r) = \sum_{\# \text{enc}} \sum_{\text{ports}} \frac{1}{V!} \left(\prod_{\text{enc}} i(\epsilon_j + \epsilon'_k) \right) \prod_{jk} \int d^2 B_{kj} e^{i(\epsilon_j + \epsilon'_k)(\frac{1}{2} + |B_{kj}|^2)} (B_{kj})^{L_{kj}^{\text{in}}} (B_{kj}^*)^{L_{kj}^{\text{out}}}. \quad (\text{C6})$$

We have thus re-expressed $Z_r^{(1)}$ as an integral over arbitrary complex $r \times r$ matrices B .

Port sum: Next, we sum over all possibilities to assign entrance and exit ports of encounters to pseudo-orbits j and k . These assignments will determine L_{kj}^{in} and L_{kj}^{out} and thus the numbers of factors B_{kj} , \tilde{B}_{jk} for the summands in Eq. (C6). To perform the sum in a compact manner, we arrange the factors B_{kj} , \tilde{B}_{jk} such that all factors originating from ports of the same encounter are grouped together. Each of these groups starts with the factor associated to the first entrance port, followed by the first exit port, the second entrance port, etc., up to the last exit port. This leads to an alternating sequence of B_{kj} 's and \tilde{B}_{jk}^* 's. As we had seen, the subscripts in this sequence must coincide as for the trace of a matrix product. We can also absorb the encounter factor from (37) into the latter sequence. The indices in that factor represent the original and partner pseudo-orbits to which the first entrance port belongs and thus must coincide with the subscripts of the first B in the sequence above. If we sum over all assignments of ports to pseudo-orbits by summing over all independent indices, each l -encounter then gives rise to a factor

$$\sum_{j_1, \dots, j_l, k_1, \dots, k_l} i(\epsilon_{j_1} + \epsilon'_{k_1}) B_{k_1, j_1} \tilde{B}_{j_1, k_2} B_{k_2, j_2} \dots B_{k_l, j_l} \tilde{B}_{j_l, k_1} = i \text{tr} \hat{\epsilon} (\tilde{B} B)^2 + i \text{tr} \hat{\epsilon}' (B \tilde{B})^2 \quad (\text{C7})$$

with

$$\begin{aligned} \hat{\epsilon} &= \text{diag}(\epsilon_1, \epsilon_2, \dots, \epsilon_r) = \text{diag}(\epsilon_A, \epsilon_C, \dots, \epsilon_C) \\ \hat{\epsilon}' &= \text{diag}(\epsilon'_1, \epsilon'_2, \dots, \epsilon'_r) = \text{diag}(\epsilon_B, \epsilon_D, \dots, \epsilon_D). \end{aligned} \quad (\text{C8})$$

If we proceed in this way, and also write the exponentials of Eq. (C6) as $\prod_{jk} e^{i(\epsilon_j + \epsilon'_k)(\frac{1}{2} + |B_{kj}|^2)} = e^{i \text{tr} \hat{\epsilon} (\frac{1}{2} + \tilde{B} B) + i \text{tr} \hat{\epsilon}' (\frac{1}{2} + B \tilde{B})}$, Eq. (C6) turns into

$$Z_r^{(1)} = \int d[B, \tilde{B}] e^{i \text{tr} \hat{\epsilon} (\frac{1}{2} + B^\dagger B) + i \text{tr} \hat{\epsilon}' (\frac{1}{2} + B B^\dagger)} \sum_{\# \text{enc}} \frac{1}{V!} \prod_{\sigma} \left(i \text{tr} \hat{\epsilon} (\tilde{B} B)^{l(\sigma)} + i \text{tr} \hat{\epsilon}' (B \tilde{B})^{l(\sigma)} \right). \quad (\text{C9})$$

Encounter sum: The remaining summation over the number V of encounters and over their sizes $l(\sigma)$, $\sigma = 1, 2, \dots, V$, is trivial. We get

$$\begin{aligned} Z^{(1)}(r) &= \int d[B, \tilde{B}] e^{i \text{tr} \hat{\epsilon} (\frac{1}{2} + B^\dagger B) + i \text{tr} \hat{\epsilon}' (\frac{1}{2} + B B^\dagger)} \sum_V \frac{1}{V!} \left(\sum_{l=2}^{\infty} \left(i \text{tr} \hat{\epsilon} (\tilde{B} B)^l + i \text{tr} \hat{\epsilon}' (B \tilde{B})^l \right) \right)^V \\ &= \int d[B, \tilde{B}] \exp \left[\frac{i}{2} \text{tr} \left(\hat{\epsilon} \frac{1 + \tilde{B} B}{1 - \tilde{B} B} + \hat{\epsilon}' \frac{1 + B \tilde{B}}{1 - B \tilde{B}} \right) \right]. \end{aligned} \quad (\text{C10})$$

2. The sigma model of RMT

Eq. (C10) agrees with the random-matrix result, which we shall derive in the following. In random-matrix theory the replica trick amounts to replacing the spectral determinants in the numerator of the generating function by powers $\Delta(E)^{-(r-1)}$ with the number of replicas r an integer larger than unity. One evaluates the thus generalized generating function for $r \geq 1$, but in the very end takes the limit $r \rightarrow 0$. This leads to the following representation of the generating function

$$\begin{aligned} Z &= \lim_{r \rightarrow 0} Z(r) \\ Z(r) &= \left\langle \Delta \left(E + \frac{\epsilon_A}{2\pi\rho} \right)^{-1} \Delta \left(E - \frac{\epsilon_B}{2\pi\rho} \right)^{-1} \Delta \left(E + \frac{\epsilon_C}{2\pi\rho} \right)^{-(r-1)} \Delta \left(E - \frac{\epsilon_D}{2\pi\rho} \right)^{-(r-1)} \right\rangle \\ &= \left\langle \prod_{j=1}^r \Delta \left(E + \frac{\epsilon_j}{2\pi\rho} \right)^{-1} \Delta \left(E - \frac{\epsilon'_k}{2\pi\rho} \right)^{-1} \right\rangle. \end{aligned} \quad (\text{C11})$$

The advantage of the replica trick is that now only inverted spectral determinants show up, which can be represented by Gaussian integrals over ordinary complex variables. If we go through the same steps as for the supersymmetric sigma model (average over matrices representing the Hamiltonian, Hubbard-Stratonovich transformation, stationary-phase approximation) we obtain the following integral representation for $Z(r)$,

$$Z(r) \sim \int_{Q^2=-1} d[Q] \exp \left(\frac{1}{2} \text{tr} \begin{pmatrix} \hat{\epsilon} & 0 \\ 0 & \hat{\epsilon}' \end{pmatrix} Q \right). \quad (\text{C12})$$

where we dropped a proportionality factor converging to 1 in the replica limit $r \rightarrow 0$. In contrast to the supersymmetric case, the matrices Q are now complex $2r \times 2r$ matrices. They have to be diagonalizable by pseudo-unitary transformations T which are now defined by

$$T \Lambda^{(1)} T^\dagger = \Lambda^{(1)}. \quad (\text{C13})$$

where $\Lambda^{(1)} = (1, 1, \dots, -1, -1, \dots)$. The saddle point manifold $Q^2 = -1$ has the shape of a hyperboloid. In the replica case the relevant solutions of $Q^2 = -1$ are matrices that can be obtained through conjugation from $i\Lambda^{(1)}$ where ,

$$Q = T \Lambda^{(1)} T^{-1}. \quad (\text{C14})$$

Multiplication with a block diagonal matrix as in the supersymmetric now leads to the rational parametrization

$$T = 1 - \begin{pmatrix} 0 & B \\ \tilde{B} & 0 \end{pmatrix} \quad (\text{C15})$$

where B is an arbitrary $r \times r$ matrix and $\tilde{B} = B^\dagger$. The fact that the off-diagonal blocks of this matrix are Hermitian-conjugate to one another follows from the pseudo-unitarity condition.

In the limit of large ϵ the integral (C12) can be calculated in a saddle point approximation (not to be mixed with the already implemented limit $N \rightarrow \infty$). Using the parametrization (C15) in Q and expanding the exponent in the vicinity of the saddle point $B = 0$ we obtain the integral

$$\begin{aligned} Z^{(1)}(r) &= \int d[B, \tilde{B}] \exp \left[i \text{tr} \hat{\epsilon} \left(\frac{1}{2} + \sum_{l=1}^{\infty} (\tilde{B}B)^l \right) + i \text{tr} \hat{\epsilon}' \left(\frac{1}{2} + \sum_{l=1}^{\infty} (B\tilde{B})^l \right) \right] \\ &= \int d[B, \tilde{B}] \exp \left[\frac{i}{2} \text{tr} \left(\hat{\epsilon} \frac{1 + \tilde{B}B}{1 - \tilde{B}B} + \hat{\epsilon}' \frac{1 + B\tilde{B}}{1 - B\tilde{B}} \right) \right] \end{aligned} \quad (\text{C16})$$

agreeing with our semiclassical formula.

Both the semiclassical and the RMT results can be generalized to the orthogonal case. One then obtains a formula as in (C16) but with B as a $2r \times 2r$ matrix subject to symmetries analogous to the supersymmetric treatment. In fact, even the crossover between the unitary and the orthogonal symmetry class is accessible in this way [44].

APPENDIX D: FERMIONIC REPLICAS

1. Semiclassical construction of a sigma model

A further variant of the sigma model makes use both replicas and Fermionic variables (the latter appearing only in the intermediate stages of the derivation from RMT). To derive this variant from semiclassics, we write the sign factor in Eq. (37) as

$$(-1)^{n_C+n_D} = (-1)^{n_A+n_B} (-1)^{(n_B+n_D)-(n_A+n_C)} \quad (D1)$$

Here the factor $(-1)^{n_A+n_B}$ can be dealt with in a similar way as the factor $(-1)^{n_C+n_D}$ in the Bosonic approach. This time we split the pseudo-orbits A and B into $r-1$ parts. There are $(r-1)^{n_A+n_B}$ ways to do this. If we take the limit $r \rightarrow 0$ this number of possibilities converges to $(-1)^{n_A+n_B}$.

The remaining sign factor $(-1)^{(n_B+n_D)-(n_A+n_C)}$ depends on the difference between the overall numbers of partner orbits and original orbits. To keep track of this factor we envisage the reconnections inside encounters as a step-by-step process. In each step the connections of just two stretches are changed, as in Fig. 5. This changes the number of orbits by one and thus flips the above sign. If we only have one 2-encounter we need only one such step and the sign factor is negative. For the example of a 4-encounter, Fig. 10 shows how we can get from connections in (a) to the final connections in (d) in three steps, each time only reshuffling two stretches. In general for each l -encounter we need $l-1$ steps. Let us denote by v_l the number of l -encounters. Then the overall number of steps, and thus sign flips, is given by the number $L = \sum_l v_l l$ of encounter stretches or links minus the number $V = \sum_l v_l$ of encounters. We are thus free to replace the sign factor by $(-1)^{L-V}$.

Altogether we then obtain

$$\begin{aligned} Z_{\text{off}}^{(1)} &= \lim_{r \rightarrow 0} Z_{\text{off}}^{(1)}(r) \\ Z_{\text{off}}^{(1)}(r) &= \sum_{\substack{\text{structures} \\ \text{of } 2r \text{ -tuples}}} \frac{1}{V!} (-1)^{L-V} \frac{\prod_{\text{enc}} i(\epsilon_j + \epsilon'_k)}{\prod_{\text{links}} (-i(\epsilon_j + \epsilon'_j))}. \end{aligned} \quad (D2)$$

Here the definition of structures is adapted to consider separately situations where the orbits in A or B are distributed differently among the $r-1$ subsets of A and B . Furthermore ϵ_j, ϵ'_k (and the corresponding diagonal matrices) are now defined as

$$\begin{aligned} \hat{\epsilon} &= \text{diag}(\epsilon_1, \epsilon_2, \dots, \epsilon_r) = \text{diag}(\epsilon_A, \dots, \epsilon_A, \epsilon_C) \\ \hat{\epsilon}' &= \text{diag}(\epsilon'_1, \epsilon'_2, \dots, \epsilon'_r) = \text{diag}(\epsilon_B, \dots, \epsilon_B, \epsilon_D). \end{aligned} \quad (D3)$$

For the diagonal term we have

$$\begin{aligned} Z_{\text{diag}}^{(1)} &= \lim_{r \rightarrow 0} Z_{\text{diag}}^{(1)}(r) \\ Z_{\text{diag}}^{(1)}(r) &= e^{i(-(r-1)\epsilon_A - (r-1)\epsilon_B - \epsilon_C - \epsilon_D)/2} (\epsilon_A + \epsilon_B)^{-(r-1)^2} (\epsilon_A + \epsilon_D)^{-(r-1)} (\epsilon_C + \epsilon_B)^{-(r-1)} (\epsilon_D + \epsilon_C)^{-1} \\ &= e^{-i \sum_j \epsilon_j / 2 - i \sum_k \epsilon'_k / 2} \prod_{j,k=1 \dots r} (-i(\epsilon_j + \epsilon'_k))^{-1}. \end{aligned} \quad (D4)$$

If we use (D2) as a starting point we can repeat the semiclassical construction of the sigma model from the previous appendix. We just have to use the new definitions of $\hat{\epsilon}, \hat{\epsilon}'$ and insert sign factors in the oscillatory prefactor, for each of the V encounters, and for each of the L encounter stretch represented by $\tilde{B}B, B\tilde{B}$. This yields

$$\begin{aligned} Z^{(1)}(r) &= \int d[B, \tilde{B}] e^{-i \text{tr} \hat{\epsilon}(\frac{1}{2} - \tilde{B}B) - i \text{tr} \hat{\epsilon}'(\frac{1}{2} - B\tilde{B})} \sum_V \frac{1}{V!} \left(- \sum_{l=2}^{\infty} \left(i \text{tr} \hat{\epsilon}(-B\tilde{B})^l + i \text{tr} \hat{\epsilon}'(-B\tilde{B})^l \right) \right)^V \\ &= \int d[B, \tilde{B}] \exp \left[-\frac{i}{2} \text{tr} \left(\hat{\epsilon} \frac{1 - \tilde{B}B}{1 + \tilde{B}B} + \hat{\epsilon}' \frac{1 - B\tilde{B}}{1 + B\tilde{B}} \right) \right]. \end{aligned} \quad (D5)$$

2. The sigma model of RMT

In RMT, the Fermionic replica sigma model is obtained if we represent *all* spectral determinants the generating function (5) by Gaussian Grassmann integrals. To that end one brings the determinants of the denominator “upstairs”

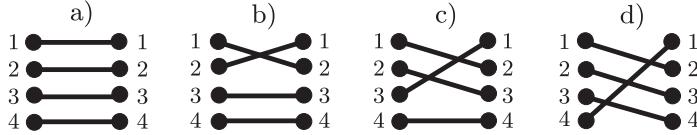


FIG. 10: Reconnections in a 4-encounter in three steps

as [34] $\det^{-1} = \det^{r-1}$, with the replica parameter r first taken as an integer larger than unity; only in the end, r is assumed real and sent to zero. The spectral determinants can be represented through Gaussian integrals over Fermionic variables. By going through the same steps as in the supersymmetric and the Bosonic replica variant, these integrals can be traded for an integral over complex $2r \times 2r$ matrices which are diagonalizable by unitary transformations and satisfy $Q^2 = -1$,

$$Z(r) \sim \int_{Q^2=-1} dQ e^{\frac{1}{2} \text{tr} Q \left(\begin{smallmatrix} \hat{\epsilon} & 0 \\ 0 & -\hat{\epsilon}' \end{smallmatrix} \right)} \quad (\text{D6})$$

with $\hat{\epsilon}, \hat{\epsilon}'$ as in (D3). The matrices Q relevant for the non-oscillatory terms in the correlator are again of the type

$$Q = iT\Lambda^{(1)}T^{-1}. \quad (\text{D7})$$

The rational parametrization then amounts to

$$T = 1 + \begin{pmatrix} 0 & -B \\ \tilde{B} & 0 \end{pmatrix} \quad (\text{D8})$$

which inserted into the integral (D6) yields

$$Z^{(1)}(r) = \int d[B, \tilde{B}] \exp \left[-\frac{i}{2} \text{tr} \left(\hat{\epsilon} \frac{1 - \tilde{B}B}{1 + \tilde{B}B} + \hat{\epsilon}' \frac{1 - B\tilde{B}}{1 + B\tilde{B}} \right) \right]. \quad (\text{D9})$$

just as in semiclassics.

Again both the semiclassical and the RMT calculation can be generalized to the orthogonal case.

The contribution of the saddle $B = \tilde{B} = 0$ does not exhaust the high- ϵ asymptotics of the generating function. Another contribution is connected with the vicinity of the so called anomalous or Andreev-Altshuler saddle, $\Lambda^{(2)} = \text{diag}(-1, 1, \dots, 1; 1, -1, \dots, -1)$ which can be obtained from $\Lambda^{(1)}$ by swapping the first and the $(r+1)$ -th entry, i. e.,

$$\Lambda^{(2)} = R\Lambda^{(1)}R^\dagger, \quad R = \begin{pmatrix} & & 1 & \\ & & \hline & 1_{r-1} \\ & 1_{r-1} & \hline & \\ 1 & & & 1_{r-1} \end{pmatrix} \quad (\text{D10})$$

where all missing entries are zero. $\Lambda^{(2)}$ cannot be reached using the rational parametrization (D8); its vicinity corresponds to B, \tilde{B} tending to infinity.

To appreciate the contribution of the Andreev-Altshuler saddle, we substitute $Q = RQ'R^\dagger$ in the integral (D6) and adopt Q' as the new integration variable. The Jacobian of this transformation is unity such that we get $Z_r = \int_{Q'^2=1} dQ' e^{\frac{1}{2} \text{tr} Q' R^\dagger X R}$ an expression still equivalent to (D6), the replacement of X by $R^\dagger X R = \text{diag}(-\epsilon'_1, \epsilon_2, \dots, \epsilon_r; \epsilon_1, -\epsilon'_2, \dots, -\epsilon'_r)$ notwithstanding. We then employ the rational parametrization for Q' and obtain the asymptotic series in ϵ^{-1} expanding the integrand around the stationary point $B = \tilde{B} = 0$ corresponding to $Q' = \Lambda^{(1)}$ or $Q = \Lambda^{(2)}$. Again, the Duistermaat-Heckmann theorem guarantees that the series breaks up after the leading term. The final result for $Z^{(2)}$ differs from $Z^{(1)}$ only by the appearance of the permuted offset matrix $R^\dagger X R$ instead of X , i. e.

$$Z^{(2)}(\epsilon_A, \epsilon_B, \epsilon_C, \epsilon_D) = Z^{(1)}(\epsilon_A, \epsilon_B, -\epsilon_D, -\epsilon_C). \quad (\text{D11})$$

The Weyl symmetry of $Z = Z^{(1)} + Z^{(2)}$ thus arises again.

APPENDIX E: STRUCTURES, PERMUTATIONS AND SYMMETRY

Here we formalize the definition of a structure in terms of permutations, in order to clarify the relation of structures to diagrams and investigate the symmetry aspects of that relation. A “diagram” is understood as the set of all physically distinct pseudo-orbit quadruplets with a given topology, i.e, with a fixed set of encounters whose ports are connected by links in a specified way. Such a diagram can have many equivalent structures associated. For simplicity we limit the following to the unitary symmetry class.

We imagine that the encounters are numbered by $\sigma = 1, \dots, V$ and the stretches and ports in the σ -th encounter by $j = 1, \dots, l(\sigma)$. Each port is thus denoted by a label $\sigma.j$. According to the convention of section IV B in the original pseudo-orbits A, C each entrance port is connected with the exit port of the same number. After the reconnection leading to the pseudo-orbits B, D the entrance port $\sigma.j$ will be connected to the exit port $\sigma.(j-1)$, with the exception of the first entrance port of each encounter which will be connected to the last exit port. The port reconnections can be depicted by a table in which the lower and the upper row list the entrance and exit ports, respectively, and each encounter is represented by a block of width $l(\sigma)$,

$$P_{\text{enc}} = \begin{pmatrix} 1.1 & \dots & 1.l(1) & 2.1 & \dots & 2.l(2) & \dots & V.1 & \dots & V.l(V) \\ 1.l(1) & \dots & 1.(l(1)-1) & 2.l(2) & \dots & 2.(l(2)-1) & \dots & V.l(V) & \dots & V.(l(V)-1) \end{pmatrix}.$$

This table can be viewed as a permutation of L numbers consisting of V independent cycles $P_{\text{enc}} = (1.l(1), \dots, 1.2, 1.1) \dots (V.l(V), \dots V.2, V.1)$, each cycle associated with an encounter. The port connections by the encounter stretches in the original pseudo-orbits A, C are described in this language by the trivial permutation I .

The connection of exit to entrance ports by links can also be represented by a similar permutation P_{link} where the upper row refers to exit ports and the lower one to entrance ports. Unlike P_{enc} the link permutation is totally arbitrary.

The permutation products $P_{\text{link}}I = P_{\text{link}}$ and $P_{\text{link}}P_{\text{enc}}$ define connections between the consecutive entrance ports before and after reconnection, respectively. Each cycle of P_{link} represents the sequence of the entrance ports of a pre-reconnection orbit in A, C ; similarly, each cycle of $P_{\text{link}}P_{\text{enc}}$ defines such a sequence for a post-reconnection orbit in B or D . The numbers n, n' of cycles in these permutations give the numbers of pre- and post-reconnection orbits.

Our definition of a structure can thus be formalized as a particular choice of (S1) the numbers V and $l(\sigma)$ determining the permutation P_{enc} , (S2) a link permutation P_{link} , and (S3) a division of the n cycles of P_{link} between the pseudo-orbits A and C , and of the n' cycles of $P_{\text{link}}P_{\text{enc}}$ between the pseudo-orbits B and D .

The order in which the encounters are numbered and the choice of the first stretch in an encounter is irrelevant. Their change amounts to port renumbering, i.e., also to a permutation (an arbitrary permutation of the encounter numbers, and a cyclic permutation of the labels for each encounter). The possible renumberings corresponding to a given $V, l(\sigma)$ form a group \mathcal{C} with $N(\mathcal{C}) = V! \prod_{\sigma=1}^V l(\sigma)$ elements. The renumbering $C \in \mathcal{C}$ transforms the encounter and link permutations as $P \rightarrow CPC^{-1}$. For a structure devoid of any symmetry, none of the non-trivial $C \in \mathcal{C}$ commute simultaneously with the link and encounter permutation, hence renumberings will produce $V! \prod_{\sigma=1}^V l_{\sigma}$ structures described by $P'_{\text{enc}} = CP_{\text{enc}}C^{-1}$, $P'_{\text{link}} = CP_{\text{link}}C^{-1}$, all of them corresponding to the same diagram and making the same contribution to the generating function. This overcounting is cancelled by the denominator $1/N(\mathcal{C})$ in (41) such that we are left with the contribution of the respective diagram in the form of the ratio of the encounter and link factors.

In a symmetric topology C may simultaneously commute with P_{link} and P_{enc} , and hence the number of equivalent structures is smaller than $N(\mathcal{C})$. However renumbering the encounters and stretches will still change the variables $s_{\sigma j}, u_{\sigma j}$ parametrizing the separations inside the encounters; the indexing of the link durations is also changed. Since we integrate over all values of $s_{\sigma j}, u_{\sigma j}$ as well as the link durations, relabelling thus leads to a separate contribution even in the symmetric case. To count each quadruplet only once we thus always have to divide by the number of possible relabellings $N(\mathcal{C})$.

Alternatively, we could formulate our approach in terms of a to sum over diagrams rather than structures. However, then the diagrams with symmetries would require special treatment. This was done in [46] for the diagrams with $L - V = 2$, but it becomes too cumbersome for a generalization to arbitrary orders.

For an example we consider a quadruplet containing, before and after the reconnection, one periodic orbit with two 2-encounters connected as shown in Fig. 11. The encounter and link permutations with the port numbering shown in Fig. 11 read

$$P_{\text{enc}} = \begin{pmatrix} 1.1 & 1.2 & 2.1 & 2.2 \\ 1.2' & 1.1' & 2.2' & 2.1' \end{pmatrix}, \quad P_{\text{link}} = \begin{pmatrix} 1.1' & 1.2' & 2.1' & 2.2' \\ 2.1 & 2.2 & 1.2 & 1.1 \end{pmatrix}$$

where we put primes over the indices of the exit ports, to distinguish them from the entrance ports. We have the following possibilities to relabel encounters and stretches (i) interchange all first indices 1 and 2, (ii) interchange 1.1

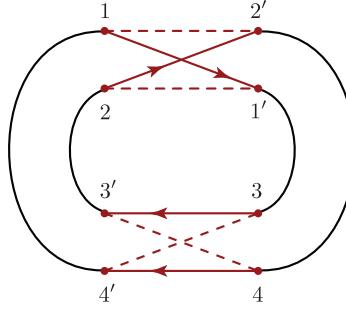


FIG. 11: A symmetric orbit pair with two encounters, already shown in the gallery of Fig. 3 as the uppermost one in the left column for $L - V = 2$; here, one of the possible port numberings is indicated.

and 1.2, and/or (iii) interchange 2.1 and 2.2. As (i) and simultaneous application of (ii) and (iii) both leave P_{enc} and P_{link} invariant there are only two equivalent structures associated with the diagram. However application of (i) and simultaneous application of (ii) and (iii) still changes the parameters $s_{\sigma j}$ and $u_{\sigma j}$. Hence when calculating the contribution of each structure by integrating over all parameters each quadruplet will be counted four times per structure. To avoid overcounting we thus have to divide out $N(\mathcal{C}) = 8$, as we would without any symmetries.

APPENDIX F: COMPLEX VS REAL CORRELATOR

Let us prove the relation $R(\epsilon) = \lim \text{Im}_{\epsilon \rightarrow +0} \text{Re } C(\epsilon)$ with the complex correlator defined by (4), in particular the presence of the constant $1/2$ in (4). Averaging in the definition of the spectral correlator, be it real or complex, is carried out over an interval of its both arguments, namely the central energy E and the energy offset $\epsilon/\pi\bar{\rho}$; to obtain a smooth function the product of the two averaging intervals must be large compared with the squared mean level spacing,

$$\delta E \delta \epsilon/\pi\bar{\rho} \gg (1/\bar{\rho})^2. \quad (\text{F1})$$

If we want oscillations of the correlator on the scale of the mean level spacing to be resolved the averaging interval of ϵ has to be taken small, $\delta\epsilon \ll 1$. In view of (F1) this means that the averaging interval $[a, b]$ of the central energy with $\delta E = b - a$ must be large compared with the mean level spacing multiplied by $1/\delta\epsilon$. On the other hand the averaging interval must be small compared with the range populated by the consecutive levels $E_1 < \dots < E_N$. (The generalization to unbounded spectra or bounded spectra with $N = \infty$ is straightforward.) We define the average over the central energy as $\langle \dots \rangle_E = (b - a)^{-1} \int_a^b dE \langle \dots \rangle$. In the energy range $E_1 \leq E \leq E_N$ the mean level spacing $1/\bar{\rho}$ can be taken as constant. Without loss of generality we choose to discuss spectral correlations in the middle of the range considered, i. e. at $E = \frac{E_N + E_1}{2}$ and can therefore take that energy as the center of the averaging interval $[a, b]$. By explicitly doing the average over E in the definition of the correlator (4) we obtain

$$C(\epsilon) + \frac{1}{2} = \left\langle \frac{1}{2\pi^2\bar{\rho}^2(b-a)} \sum_{i,k=1}^N \frac{1}{\frac{\epsilon^+}{\pi\bar{\rho}} + E_i - E_k} \left[\ln \frac{x - \frac{\epsilon^+}{2\pi\bar{\rho}} - E_i}{x + \frac{\epsilon^+}{2\pi\bar{\rho}} - E_k} \right]_{x=a}^{x=b} \right\rangle_{\epsilon}.$$

where the remaining average is over ϵ only. Using $(x + i\eta)^{-1} \leftrightarrow \mathcal{P} x^{-1} - i\pi\delta(x)$ for $\eta \rightarrow 0^+$ and $\ln x = \ln|x| + i\text{Arg } x$, we can write the real part of C as

$$\text{Re } C(\epsilon) = -\frac{1}{2} + R_{I,I} + R_{R,R} \quad (\text{F2})$$

where $R_{I,I}$ ($R_{R,R}$) stem from products of the imaginary (real) parts of the fraction and the square bracket.

We begin with $R_{I,I}$ assuming $\text{Im } \epsilon^+ = \eta \rightarrow +0$. In the summand (i, k) the phase of the expression in the square brackets can be (i) 2π (both E_i and E_k lie within the averaging interval of the central energy $[a, b]$); (ii) π (only one eigenvalue lies within $[a, b]$); (iii) zero (both eigenvalues are outside $[a, b]$). Let us assume that the energy offset $\epsilon/\pi\bar{\rho}$ is small compared with $b - a$; then considering the factor $\delta \left(\frac{\epsilon}{\pi\bar{\rho}} + E_i - E_k \right)$, we may neglect the possibility (ii). Up

to an additive constant, $R_{I,I}$ must then coincide with the real correlation function,

$$R_{I,I} = \frac{1}{(b-a)\bar{\rho}^2} \left\langle \sum_{a < E_i, E_k < b} \delta \left(\frac{\epsilon}{\pi\bar{\rho}} + E_i - E_k \right) \right\rangle_{\epsilon} = R(\epsilon) + 1.$$

In the sum of products of real parts

$$R_{R,R} = \frac{1}{2\pi^2 \bar{\rho}^2 (b-a)} \left\langle \sum_{i,k=1}^N \frac{1}{\frac{\epsilon}{\pi\bar{\rho}} + E_i - E_k} \ln \left| \frac{(b - \frac{\epsilon}{2\pi\bar{\rho}} - E_i)(a + \frac{\epsilon}{2\pi\bar{\rho}} - E_k)}{(a - \frac{\epsilon}{2\pi\bar{\rho}} - E_i)(b + \frac{\epsilon}{2\pi\bar{\rho}} - E_k)} \right| \right\rangle_{\epsilon} \quad (\text{F3})$$

the poles of $\left(\frac{\epsilon}{\pi\bar{\rho}} + E_i - E_k \right)^{-1}$ are cancelled by the zero of the respective logarithms; the mild logarithmic singularities of $R_{R,R}$ at $E_i = a, b \pm \frac{\epsilon}{2\pi\bar{\rho}}$ are unconnected with the level correlation and thus of no physical relevance. (They are caused by the rectangular cut-off of the integrand at the borders of the averaging interval in $\langle \dots \rangle_E$ and disappear if this cut-off is smoothed out). Hence the summand has no relevant singularities or other fluctuations on the scale of ϵ . We are thus free to replace the double sum over a discrete eigenvalue spectrum by a double integral over a uniformly distributed density, as in

$$R_{R,R} = \frac{1}{2\pi^2 (b-a)} \int_{E_1}^{E_N} dx dy \frac{1}{\left(\frac{\epsilon}{\pi\bar{\rho}} + x - y \right)} \ln \left| \frac{(b - \frac{\epsilon}{2\pi\bar{\rho}} - x)(a + \frac{\epsilon}{2\pi\bar{\rho}} - y)}{(a - \frac{\epsilon}{2\pi\bar{\rho}} - x)(b + \frac{\epsilon}{2\pi\bar{\rho}} - y)} \right|.$$

Note that the smoothing w.r.t. the offset variable is thus rendered superfluous.

Consider now the semiclassical limit $\hbar \rightarrow \infty$ which means that $\bar{\rho} \rightarrow \infty$; we also assume the spectral span constant (in classical units) which means that the number N of levels taken into account goes to infinity. Let the averaging interval $[a, b]$ simultaneously shrink compared with the spectral span in such a way that the hierarchy $E_N - E_1 \gg b-a \gg \epsilon/\pi\bar{\rho}$ is observed. Neglecting $\frac{\epsilon}{\pi\bar{\rho}}$ and letting the parameter $A \equiv \frac{E_N - E_1}{b-a}$ go to infinity we get

$$\begin{aligned} R_{R,R} &= \frac{1}{4\pi^2} \lim_{A \rightarrow \infty} \int_{-A}^A du dv \frac{1}{u-v} \ln \left| \frac{(1-u)(1+v)}{(1+u)(1-v)} \right| \\ &= -\frac{2}{\pi^2} \int_0^1 \frac{dx}{x} \ln \left| \frac{1+x}{1-x} \right| = -\frac{1}{2}. \end{aligned} \quad (\text{F4})$$

Collecting (F2), (F3), and (F4) we come to the desired relation, $\lim_{\eta \rightarrow 0} \text{Re } C(\epsilon^+) = R(\epsilon)$.

- [1] O. Bohigas, M. J. Giannoni, and C. Schmit, Phys. Rev. Lett. **52**, 1 (1984); S. W. McDonald and A. N. Kauffmann, Phys. Rev. Lett. **42**, 1189 (1979); G. Casati, F. Valz-Gris, and I. Guarneri, Lett. Nuovo Cim. **28**, 279 (1980); M. V. Berry, Proc. R. Soc. Lond. A **413**, 183 (1987).
- [2] E. P. Wigner, Proc. 4th Can. Math. Congr., Toronto, 174 (1959).
- [3] F. Haake, *Quantum Signatures of Chaos*, 2nd ed. (Springer, Berlin, 2001).
- [4] K. Efetov, *Supersymmetry in Disorder and Chaos* (Cambridge University Press, Cambridge, England, 1997).
- [5] T. Guhr, A. Müller-Groeling, and H.A. Weidenmüller Phys. Rep., **299**, 189 (1998)
- [6] A.D. Sokal, J. Math. Phys. **21**, 261 (1980).
- [7] M. C. Gutzwiller, *Chaos in Classical and Quantum Mechanics* (Springer, New York, 1990).
- [8] N. Argaman et al., Phys. Rev. Lett. **71**, 4326 (1993).
- [9] M. V. Berry, Proc. R. Soc. Lond. A **400**, 229 (1985).
- [10] A. Altland et al., *Near action-degenerate periodic-orbit bunches: a skeleton of chaos*, in Path Integrals—New Trends and Perspectives, Proc. of 9th Int Conference, W. Janke and A. Pelster, eds., World Scientific, 40 (2008)
- [11] I. L. Aleiner and A. I. Larkin, Phys. Rev. B **54**, 14423 (1996); R. A. Smith, I. V. Lerner, and B. L. Altshuler, Phys. Rev. B **58**, 10343 (1998); R. S. Whitney, I. V. Lerner, and R. A. Smith, Waves Random Media **9**, 179 (1999).
- [12] M. Sieber and K. Richter, Physica Scripta **T90**, 128 (2001); M. Sieber, J. Phys. A **35**, L613 (2002).
- [13] S. Müller, S. Heusler, P. Braun, F. Haake, and A. Altland, Phys. Rev. Lett. **93**, 014103 (2004).
- [14] S. Müller, S. Heusler, P. Braun, F. Haake, and A. Altland, Phys. Rev. E **72**, 046207 (2005).
- [15] D. Waltner, S. Heusler, L.D. Urbina and K. Richter, J. Phys. A, **42**, 292001 (2009).
- [16] J. P. Keating, *The Riemann zeta-function and quantum chaology*, in Quantum Chaos, G. Casati, I. Guarneri, and V. Smilansky, eds., North-Holland, Amsterdam, 145 (1993).

- [17] E. B. Bogomolny and J. P. Keating, Phys. Rev. Lett. **77**, 1472 (1996).
- [18] S. Heusler, S. Müller, A. Altland, P. Braun, and F. Haake, Phys. Rev. Lett. **98**, 044103 (2007).
- [19] M. V. Berry, *Riemann's zeta function: a model for quantum chaos?* In Quantum chaos and statistical nuclear physics (eds T. H. Seligman & H. Nishioka). Springer Lecture Notes in Physics, no. 263, Berlin, Germany: Springer (1986); M.V. Berry and J.P. Keating, J. Phys. A **23**, 4839 (1990); J.P. Keating, Proc. R. Soc. Lond. A **436**, 99 (1992); M.V Berry and J.P. Keating, Proc. R. Soc. Lond. A **437**, 151 (1992).
- [20] J. P. Keating and S. Müller, Proc. R. Soc. A **463**, 3241 (2007).
- [21] M. Zirnbauer, *Generating function for GUE, GOE and crossover cases*, private communication (2008)
- [22] B. A. Muzykantskii and D. E. Khmel'nitskii, JETP Lett. **62**, 76 (1995); A. V. Andreev, O. Agam, B. D. Simons, and B. L. Altshuler, Phys. Rev. Lett. **76**, 3947 (1996).
- [23] J. Müller and A. Altland, J. Phys. A **38**, 3097 (2005); J. Müller, T. Micklitz, and A. Altland, Phys. Rev. E **76**, 056204 (2007)
- [24] J. H. Hannay and A. M. Ozorio de Almeida, J. Phys. A **17**, 3429 (1984).
- [25] M. Turek, D. Spehner, S. Müller, and K. Richter, Phys. Rev. E **71**, 016210 (2005).
- [26] P. A. Braun, F. Haake, and S. Heusler, J. Phys. A **35**, 1381 (2002).
- [27] D. Spehner, J. Phys. A **36**, 7269 (2003).
- [28] M. Turek and K. Richter, J. Phys. A **36**, L455 (2003).
- [29] S. Müller, Eur. Phys. J. B **34**, 305 (2003).
- [30] R. Bowen, Amer. J. Math. **94**, 1 (1972); W. Parry and M. Pollicott, Astérisque **187**, 1 (1990).
- [31] S. Müller, PhD thesis, Universität Duisburg-Essen, [nlin.CD/0512058](#) (2005).
- [32] S. Heusler, S. Müller, P. Braun, and F. Haake, Phys. Rev. Lett **96**, 066804 (2006); P. Braun, S. Heusler, S. Müller, and F. Haake, J. Phys. A: Math. Gen. **39**, L159 (2007); S. Müller, S. Heusler, P. Braun, and F. Haake, New J. Phys. **9**, 12 (2007).
- [33] I. V. Yurkevich and I. V. Lerner, Phys. Rev. B **60**, 3955 (1999).
- [34] A. Kamenev and M. Mézard, J. Phys. A: Math. Gen. **32**, 4373 (1999).
- [35] A. V. Andreev and B. L. Altshuler, Phys. Rev. Lett. **75**, 902 (1995)
- [36] M.L. Mehta, *Random matrices* (3rd ed., Elsevier/Academic Press, Amsterdam, 2004).
- [37] H.-J. Stöckmann, *Quantum Chaos - An Introduction* (Cambridge Univ. Press, Cambridge, 1999)
- [38] P. Brouwer and A. Altland, Phys. Rev. B **78**, 075304 (2008).
- [39] P. Brouwer and S. Rahav, Phys. Rev. Lett. **96**, 196804 (2006); R. S. Whitney and P. Jacquod, Phys. Rev. Lett. **96**, 206804 (2006).
- [40] E. B. Bogomolny, B. Georgeot, M. J. Giannoni, and C. Schmit, Phys. Rev. Lett. **69**, 1477 (1992).
- [41] T. Nagao and S. Müller, J. Phys. A: Math. Theor., **42**, 375102 (2009).
- [42] K. Richter and M. Sieber, Phys. Rev. Lett., **89**, 206801 (2002)
- [43] G. Berkolaiko, J. M. Harrison, and M. Novaes, J. Phys. A: Math. Theor. **41** 365102 (2008).
- [44] K. Saito and T. Nagao, Phys. Lett. A **352**, 380 (2006); T. Nagao, P. Braun, S. Müller, K. Saito, S. Heusler, and F. Haake, J. Phys. A: Math. Theor. **40**, 47 (2007); J. Kuipers and M. Sieber, J. Phys. A: Math. Theor. **40**, 935 (2007); K. Saito, T. Nagao, S. Müller and P. Braun, Preprint [arXiv:0906.2033v1](#) ([nlin.CD](#)), (2009).
- [45] J. J. Duistermaat, G. J. Heckman, Invent. Math. **69**, 259 (1982)
- [46] S. Heusler, S. Müller, P. Braun and F. Haake, J. Phys. A: Math. Gen. **37**, L31 (2004)